Self-Expression

Mitchell S. Green

Self-Expression

Mitchell S. Green presents a systematic philosophical study of self-expression—a pervasive phenomenon of the everyday life of humans and other species, which has received scant attention in its own right. He explores the ways in which self-expression reveals our states of thought, feeling, and experience, and he defends striking new theses concerning a wide range of fascinating topics: our ability to perceive emotion in others, artistic expression, empathy, expressive language, meaning, facial expression, and speech acts. He draws on insights from evolutionary game theory, ethology, the philosophy of language, social psychology, pragmatics, aesthetics, and neuroscience to present a stimulating and accessible interdisciplinary work.

This page intentionally left blank

Self-Expression

Mitchell S. Green
University of Virginia

For Lori

This page intentionally left blank

Acknowledgments

Research for this book was greatly facilitated by generous institutional support, most importantly a Burkhardt Fellowship for Recently Tenured Scholars, offered by the American Council of Learned Societies. I held that fellowship at the National Humanities Center, which provided a friendly and stimulating environment in academic year 2001–2 that enabled me to do the heavy conceptual lifting that formed the core themes of this book. Later research support came from the Center for Contemplative Mind in Society, as well as from the Teaching Resource Center at the University of Virginia and the Office of the Provost at the University of Virginia.

Many people worked selflessly to offer comments on earlier drafts. Ana Balan, Dorit Bar-On, Wayne Davis, Chris Gauker, and Garry Hagberg went above and beyond the call of duty to comment on the entire draft of the book, and I learned a great deal from their insights. Two anonymous referees for Oxford University Press also offered detailed and insightful comments, and I thank them for their contributions. Others offered comments on one or more element of the developing project: Kent Bach, Ayca Boylu, Cynthia Freeland, Bernard Gert, Michael Kubovy Marc Lange, Jerry Levinson, Bill Lycan, Al Martinich, Joseph Moore, Ram Neta, Jenefer Robinson, Keith Simmons, Robert Stecker, and Ken Walton all offered insights, and I am grateful for their constructive criticisms. In addition, Peter Momtchiloff, Catherine Berry, and Victoria Patton of Oxford University Press provided judicious guidance and encouragement, and I could not have asked for better editorial support.

Bits of this project have been presented at the following institutions: The Central Division of the American Philosophical Association, The American University in Beirut, Amherst College, Central Michigan University, Gesellschaft für Analytische Philosophie, The International Pragmatics Association, the National University of Singapore, Northwestern University, Singapore Management University, Texas A&M University, the Virginia Philosophical Association, Virginia Commonwealth University, Virginia Tech, the University of North Carolina, the University of Virginia, and

Western Michigan University. I thank receptive and stimulating audiences at these institutions for their criticism and encouragement.

Finally, I thank my family, Lori, Noah, Sofia, and Stella, for their unending support, patience, encouragement, and love. Without them it would not have been possible to write this book.

Contents

Figures and Tables

For all men live by truth, and stand in need of expression. In love, in art, in avarice, in politics, in labor, in games, we study to utter our painful secret. The man is only half himself, the other half is his expression.

R.W. Emerson, 'The Poet'

1

The Significance
of Self-Expression

We express ourselves in many ways: through tone of voice, posture, the face, words, and, in more subtle cases, paint, music, sculpture, and other forms of art. Linking these disparate phenomena together is a pattern of behavior coping with a felt need. It is one thing to bring someone's attention to the prey, the water source, the lurking danger, or your injury. It is another thing to make them aware of how things appear from your point of view. Your point of view, including beliefs, emotions, moods, and experiences, can't be made manifest in the way that you can make this bruise or that snake manifest. Yet for eliciting another's aid, succor, sex, or submission, nothing beats an emotional display; for coordinating action a display of belief or intention is virtually indispensable. Philosophers, whose business it is to question platitudes, rightly wonder how it's possible to know what's in someone else's heart or mind. However, their current answers to this question underestimate just how ingenious our and other species have been in solving the problem of manifesting one's point of view. Self-expression is the measure of that ingenuity.

In expressing ourselves we manifest some part of our point of view. Given the obsession that students of the mind have had with one or another aspect of points of view—including beliefs, emotions, moods, and experiences—you might expect to find self-expression at the top of a high pedestal within the cognitive sciences. I, for one, have found it nearly everywhere else. Consider a cell-phone ad, an art class syllabus, a commentary on a bit of music, a clinical description of a child's behavior: In each of these one would not be surprised to find the notion of self-expression. The ad might enjoin you to express yourself, the syllabus might promise to teach you how to do so, the commentary might speak of how the artist is eloquently expressing herself, and the clinical description

might speak of the child's difficulty in doing so. Talk of self-expression is pervasive in everyday life. Each of these examples even has to do with communication—over the phone, through paint or music, in the child's verbal and non-verbal behavior. Yet as a topic of research, self-expression is on the margins of both contemporary linguistics and that part of contemporary philosophy concerned with communication, the philosophy of language. Further, where it is often invoked, in social psychology, it is little scrutinized. As a result, self-expression is a pervasive theme of everyday life little dignified by the attention of cognitive scientists.

They have, admittedly, had other things to do. The last century of the philosophy of language, and the last four decades of linguistics, have seen explosions of research in the fields of syntax and semantics. For many purposes, researchers in these fields can justifiably ignore expressive dimensions of communication in order to produce recursive, compositional characterizations of syntactic and semantic phenomena abstracted away from their use to express thought, feeling, or experience. Such characterizations have revolutionized our understanding of language and communication. Yet while linguists and philosophers of language do not deny that language has expressive components, these components have no natural home in their theories.

Things were different in the nineteenth century and the early part of the twentieth. Much philosophy in that period, and a good deal of the psychology, literary theory, and musicology that grew out of it, gave pride of place to self-expression. The artist, it was said, expresses himself in his work, enabling the audience, if she be both sensitive and brave, to commune with him. On many versions of this picture, the artist need not be human, and the medium of communion need not be, or even be translatable into, words. Thus Nathaniel Hawthorne, after a day of wandering through the Italian countryside, writes in his journal:

myself, with J————for my companion, kept on even to the city gate,—a distance, I should think, of two or three miles at least. The lower part of the road was on the edge of the hill, with a narrow valley on our left; and as the sun had now broken out, its verdure and fertility, its foliage and cultivation, shone forth in miraculous beauty, as green as England, as bright as only Italy. Perugia appeared above us, crowning a mighty hill, the most picturesque of cities; and the higher we ascended, the more the view opened before us, as we looked back on the course that we had traversed, and saw the wide valley, sweeping down and spreading out, bounded afar by mountains, and sleeping in sun and shadow. No language nor any

art of the pencil can give an idea of the scene. When God expressed himself in the landscape to mankind, He did not intend that it should be translated into any tongue save his own immediate one. (Hawthorne 1900)

If art expresses what's within rather than representing what's without, we needn't worry that the fast-improving technology of photography, and then, soon upon its heels, moving pictures, can depict the world as well as Vermeer or Velasquez. A Romantic theory of art enables us to deny that this was their point anyway. Art is not so much mirror as lamp.

Philosophical interest in self-expression in the latter part of the nine-teenth century tended to be associated with controversial theories of art, knowledge, meaning, or emotion. According to some views, all art is a form of self-expression; for others, all self-expression is also an intuition of a bit of knowledge that is not translatable into words or any other medium; yet others hold that what is expressed does not exist until the artist's act of expression. One can forgive positivists of the early twentieth century for professing to find such positions either gibberish or false. Whether or not this reaction was justified, however, self-expression seems, by the middle of that century, to have been found guilty by association with dubious philosophy. Its reputation was hardly burnished by comporting with ethics. So-called expressivist theories of ethics took such words as 'good' and 'bad' as not being in the business of ascribing properties to the world but as expressing attitudes (approbation, denigration, etc.) of the speaker. While they still have their defenders today, such views are by no means in the mainstream.

The notion of self-expression has been collecting dust for some decades now. Not so with expressiveness. An artist might create a work that is express*ive* without expressing his or her own feelings; Beethoven is said to have brought audiences to tears while smirking through his perfor-mances. Accordingly, in the last few decades aestheticians, after dissociating expressiveness from its shady friends, have persistently studied it in music, painting, literature, and other of the arts. This tradition has produced much innovative work, some of which we'll be considering in Chapter 7. We'll see, however, that aestheticians largely explain expressiveness in terms of self-expression, taking the latter notion for granted. While that doesn't vitiate their theories, I'll try to show how a fuller account of self-expression enables us to make progress on some current problems in aesthetics.

I've no wish to put self-expression, or any related notion, atop a pedestal. Rather, in this first book-length study of expression since Tormey 1971, I will approach the topic as a key to understanding how we share our point of view. This approach will shed light not only on problems within aesthetics, but also on issues in the philosophy of mind and language, linguistics, experimental psychology, and even the evolution of language. Because, as I've already suggested, self-expression is bound up with communication, I'll start by situating it within accounts of this phenomenon.

1.1. Four models of communication

Pragmatics, once the impoverished neighbor of syntax and semantics, has recently come into its own and is now the locus of active empirical and theoretical investigation. Associated with major trends in pragmatic theory are distinctive pictures of communication that promise to encompass syntax and semantics, but much more as well. It might be hoped that one or more such picture and its associated pragmatic doctrine would offer a foster home to orphan expression. However, after considering two popular models of communication bound up with movements in pragmatics—the code model and the inferential model—I'll argue that self-expression is suited to neither. In addition, a third pragmatic model of communication, which I'll call the extended senses model and which has recently been winning adherents, requires modification if it is to accommodate self-expression. That modification is inspired by a very different approach to communication that can be found in both computer science and evolutionary biology—what I'll call the signaling model. Subsequent chapters will articulate this synthesis of the extended senses and signaling models as well as the notion of self-expression that suggests it. This articulation will in turn shed new light on many other topics including speaker meaning, speech acts, knowledge of other minds, implicature, the psychology and evolutionary biology of facial expression, empathy, qualia, and artistic expression.

What, then, is communication? This notion is used in both a permissive and a restrictive sense. In the permissive sense, such as one finds in contemporary evolutionary biology, ethology, or computer science, communication is a matter of successful signaling. I will clarify this with

some definitions. A *cue* is any feature of an entity that conveys information (including misinformation). That information might pertain to how things were, how things are, how things will be, or how things ought to be. A tiger's pug marks cue other animals to the presence of a predator; my emission of CO_2 cues mosquitoes to the presence of a food source. In neither of these cases need the information be something that the organism aimed, planned, or was designed to convey. The tiger and I would prefer to keep our whereabouts, and our blood, to ourselves.

A *signal* is any cue that was designed for its ability to convey the information that it does.[1] The design in question might be due to the work of an intelligent agent or agents, or be the product of evolution by either artificial or natural selection. Products of intelligent design include computers that communicate with one another across a network. Products of evolution by artificial selection include animals that have been bred for certain traits. Products of evolution by natural selection include most of the biological species known to us. Bright coloration in an anuran, for instance, signals potential predators that it's either unpleasant tasting or noxious if eaten. When the process of natural selection accounts for a trait as an adaptation, that shows it to be designed without thereby justifying the inference that there must have been an intelligent design*er*.

Some signals might be due in part to evolution by natural selection and in part to the choice of an intelligent, conscious agent. Many features of natural language presumably result from evolution by natural selection, and if that is so then a human being's choice of a sentence to convey a thought is the product of both evolution and intelligent agency. When a signal succeeds in conveying the information for which it was designed, communication takes place.[2] Call this the *signaling model of communication*.

---

[1] This definition of a signal is close to the notion of communication found in Johnson-Laird 1991, and may be seen as a generalization of a notion offered in recent evolutionary biology. Maynard Smith and Harper offer the following definition: ''We define a 'signal' as any act or structure which alters the behaviour of other organisms, which evolved because of that effect, and which is effective because the receiver's response has also evolved.'' (Maynard Smith and Harper, 2004, p. 3.) I shall instead call this a *biological signal*, leaving room for a more general notion that applies to things other than organisms and to processes other than evolution. Also, their usage makes it clear that Maynard Smith and Harper mean a signal to be something that *tends* to alter the behavior of other organisms, so that a creature can send a signal that no other organism receives. I shall follow them in this.

[2] Perhaps architects have an even more permissive use than this, as exemplified by such remarks that this kitchen communicates with that foyer. Here we have design, but what is communicated is not information but inhabitants.

It is not true by definition that the information conveyed by a signal is accurate; signals can also contain misinformation, and many animals exploit this fact to their advantage. For instance, to escape predation some anurans sport bright colors even when they are neither poisonous nor noxious to predators. However, although in any given case a signal can misrepresent a state of affairs and may even deceive its receiver, it is important to the stability of any signaling system that it be on the whole reliable. If the only thing that ever glowed deep under the sea were the lure of the angler-fish, he'd be out of business after a while. The fishes that tend to pursue his deceptive lure would eventually be eaten and thus stop contributing genes to their species. Natural selection, and as I shall try to show in Chapters 3 and 4, also culture, tend for this reason to settle upon ways of vouchsafing the veracity of signals. Among their ways of doing so are signals that can only be faked with great difficulty due to limitations on the organism.

Here is an example: A tiger reaches as high as it can on the bark of a tree to make scratches marking its territory (Thapar 1986). The height of the scratch is thus a reliable indicator of the size of the tiger. Its size is in turn a good indicator of its ability to defend its territory.[3] The height of the scratch marks, indeed, shows not just the tiger's size, but also its ability to defend its territory. Further, it is plausible that scratching as high as it can also signals that ability. (Whether or not this is so is an empirical question to be settled by further ethological and evolutionary research.) If it is a signal, then it is one that is very difficult for the tiger to fake: One can envision the tiger jumping off a nearby branch, scratching high on a tree trunk before alighting on the ground, and so forth, but such cases are pretty far-fetched. Consequently, the chances of a deceptive such signal are very small. Signals that can only be faked with great difficulty as a result of limitations on the organism are known as *indices*.

Another way of ensuring the difficulty of faking a signal is by making it costly to produce, specifically, more costly than is required just to produce a signal of that type. Male peacocks have flamboyant arrangements of feathers making them less agile and easier for predators to spot; growing such feathers also costs extra calories (Zahavi and Zahavi 1997). Such feathers nevertheless give males an advantage in sexual selection (Petrie,

---

[3] Biologists generalize such notions as an animal's ability to defend its territory with the concept of resource holding potential (RHP), but we won't need to do that here.

Halliday, and Sanders 1991). It is not just that ostentatious feathers make a male easier for the female to spot. Rather, an ostentatious display is like saying, "Just think of how fit I must be if I can survive with this baggage!" Signals that can only be faked with great difficulty as a result of being very costly to produce are known as *handicaps*. A signal's being costly to produce is thus one way of being difficult to fake because of limitations on an organism. As a result, a handicap is a special case of an index.[4]

Many researchers in the cognitive sciences tolerate the permissive notion of communication, and would probably have no objection to the signaling model thereof including its subcategories such as indices and handicaps. However, such researchers tend to distinguish that permissive notion from a more restrictive one that seems closer to their concerns. This more restrictive notion of communication focuses exclusively on the conveyance of information about cognitive or affective states such as belief, desire, intention, and various emotions. One such cognitive model of communication is deeply rooted in our thought, both literal and metaphorical. On this model, when communication takes place a communicator C encodes her ideas (or thoughts, beliefs, judgments, emotions, or other features of her state of mind or heart) into a medium, and sends them to addressee A for decoding. If all goes well, A will be able to unpack and decode the parcel, and on that basis gain information from C. For members of our species the code in question will typically be a natural language with its conventional meanings, but it could also be an artificial language such as semaphore or first order predicate logic. Either way, at the heart of communication is the coding and conveyance of information about cognitive states, with the aim of producing cognitive effects such as belief or knowledge in addressees. Following Sperber and Wilson 1995, call this the *code model of communication*.

Adopting a more restrictive use of the notion of communication is a terminological choice that in itself is innocuous enough. Even accepting that choice, however, one might take issue with the code model's reliance

---

[4] Maynard Smith and Harper do not define handicaps as species of index; instead they appear to treat handicaps and indices as two incompatible ways of being a signal. However, given their definition of an index ("a signal whose intensity is causally related to the quality being signaled, and which cannot be faked"), it seems to me to follow that a signal whose cost ensures its reliability will be an index. After all, such a signal will be causally related to the quality being signaled: For instance, if the peacock were not viable, he would not have been able to survive with that exotic train. For these reasons, I will treat handicaps as a species of index.

on the notion of a code. For instance, we communicate much more than we literally say. I ask you where John is and you respond, ''He's either in the kitchen or the garage.'' You most likely communicate not only that John is either in the kitchen or the garage, but that you don't know which it is. Yet this latter bit—that you don't know in which of these two places John is to be found—is not communicated by virtue of a code. Rather, I figure out that you must mean some such thing as this, since if you had known more you would have said more. While codes, such as one finds in the conventional features of natural languages, are certainly a pervasive part of communication, it is doubtful that they are essential for communication, and thus doubtful that they should be essential to a model thereof.

The limitations of the code model suggest another cognitive model that, while perhaps not as entrenched in our self-conscious theorizing about language as is the code model, is nevertheless quickly supported by reflection on everyday experience. On this view, when communication takes place a communicator C intentionally provides evidence for some addressee A about C's cognitive state. A's job is then, on the basis of that evidence, to draw inferences about C's cognitive state. That inference will either be carried out self-consciously, or be open to conscious introspection should A care to reflect on her own ability to discern what C is trying to convey. If A draws a correct, or nearly correct, inference about C's cognitive state, then C and A have communicated; otherwise C has attempted but failed to communicate with A. C's success in communicating, on this model, requires consideration of what inferences A will be able to perform in discerning C's message. Again following Sperber and Wilson, we may call this the *inferential model* of communication. The inferential model makes sense of some cases that the code model seems unable to accommodate. For instance, it makes sense of how an addressee can discern how a communicator conveys more than she literally says, such as in the example we just considered concerning John's whereabouts.[5]

The code and inferential models of communication are consistent with one another in that one might be used to account for one range of

---

[5] An earlier and influential version of the inferential model is Bach and Harnish 1979. Sperber and Wilson (1995, ch. 2) make clear that inference involves deduction carried out on syntactic objects, and may be spontaneous as well as unconscious. However, they are also clear that such inference is nevertheless *open* to conscious introspection. Thus, for instance, most visual perception is not an inferential process, while much pragmatic comprehension is.

phenomena while the other is used to account for another without the theorist who uses them in this way contradicting herself. On the other hand, parsimony might tempt one to promote one of these two models as being fundamental, perhaps construing the other as a special case. We have already seen that the code model is not general enough to play the former role; might the inferential model be instead? According to this suggestion, addressees of communication always perform inferences to access speakers' meanings; in some cases, however, for instance when natural or artificial language is used, they also need to invoke the meaning of a code in order to do so. Yet this suggestion is not tenable. The reason is that when we speak to one another using a natural language in which we are fluent, we often understand one another without recourse to inferential processes—where, as before, an inferential process is defined as either conscious or open to conscious introspection. In the most typical cases, when I hear your utterance of 'It's raining', I know what you mean, but cannot introspect on the process by which I turn the sounds I hear into an interpreted message. Instead, our understanding of one another in cases such as these is instead typically automatic, and neither conscious nor open to conscious scrutiny.

Ecumenicism may seem the wisest route at this point. Perhaps the code model should be retained to make sense of some communicative phenomena, while the inferential model should be retained to make sense of others; we might then go in search of another model to cover any cases that might remain. Before taking this path, however, we do well to observe that both the code and inferential models, by restricting their focus to communication of states of mind, suggest that this is the *point* of communication, at least among species such as our own. That suggestion might be challenged. After all, I typically am interested in your state of mind as a route to finding out about the world: if I have no faith that what you think bears on how things are, I am likely to lose interest in what you think except for the special case in which I am concerned to find out how you are going to act.

This observation suggests another model of communication that foregrounds the use we make of one another as sources of information about the world around us. According to what I shall call the *extended senses* model of communication, the primary aim of communicating is not the conveyance of information about the intentional states of communicators.

Rather, our primary aim in communicating with one another is to widen each other's perceptual reach. Thus Michael Dummett: "we learn to react to the statements of others in the same way that we react to various observed features in the environment… It is thus essential to the activity of assertion that the making of an assertion will in general modify the behaviour of those to whom it is made" (1981, p. 355). And John McDowell: "The primary point of asking questions is not to acquire beliefs about one's interlocutor's beliefs, but to find out how things are. Correspondingly, the primary point of making assertions is not to instil into others beliefs about one's own beliefs, but to inform others—to let them know—about the subject matter of one's assertions" (1980, p. 38). More recently, Ruth Millikan has espoused an even stronger position: "Forming a belief about where Johnny is on the basis of being told where he is, is just as direct a process (and just as indirect) as forming a belief about where Johnny is on the basis of seeing him there" (2004, ch 9, p. 120). Since forming a belief about Johnny's presence on the basis of seeing him there does not involve inference, Millikan is in effect setting her face against the inferential model of communication.

The extended senses model of communication suggests that each member of a communicative group gains, by virtue of that membership, vastly enhanced perceptual powers. It is as if, for each member of such a group, all the other members' senses are prostheses. The model is right thus to stress our intimate informational dependence upon one another: it's a good bet that our own species' survival has been enhanced greatly by this pooling of perceptual resources. It is, further, not difficult to think of other species of which much the same may be said. Similarly for interspecific communication: Human–canine co-evolution in effect provides improved hearing for us, and improved eyesight for the canids. To its credit, the extended senses model doesn't require us to choose the more restrictive sense of 'communication' over the more permissive sense of that term. The extended senses model is consistent with the idea that some signals convey information in a way that is the product of design irrespective of anyone's intentions or plans. A cognitive scientist has no reason to oppose it just by virtue of her institutional affiliation.

In spite of these virtues, the extended senses model of communication overreacts to the code and inferential models' undue focus on the mental state of the communicator. It does so by failing to tell us what, if anything,

is important about that dimension of communication. Surely, just as we depend upon others for information about the world, so too we depend on their attitudes toward that information: not only is it of interest to learn from my colleague that something large is behind that tree, the abject terror on her face communicates an appropriate response—my own fear, or at least my vigilance. Again, the extended senses model treats us as all on a par with respect to one another: we are all pretty much equal with respect to our ability to serve as prostheses for one another's senses. However, the extended senses model pays inadequate attention to the fact that as members of a social group we are forever guided by relations of dominance and submission, alliances, negotiations, détente, and the like. A conspecific's gaze might indicate the presence of food, but the threatening look on his face indicates the danger of my trying to get it. The dog's bark indicates the presence of an unfamiliar creature, and the urgency of that bark indicates the need for an immediate response.

While one might accept the extended senses model without committing oneself to Millikan's strong formulation of that position, we also do well to tread carefully as we consider different ways of forming beliefs about Johnny's whereabouts. After all, if I form such a belief by seeing where he is, I gain perceptual knowledge (of what Johnny looks like, how he is standing, etc.) that I don't get just by being told where he is. And yet recall the tiger: its scratch marks show its size, and thus show one aspect of what it looks like. Similarly, expressive dimensions of communication can telegraph perceptual as well as affective knowledge: the look on your face, your tone of voice, and so on as you convey Johnny's whereabouts might convey to me whether the situation you report is cause for fear, delight, or relief. May we accommodate these dimensions of communication in a more comprehensive model?

1.2. Signals and expressions

I believe the answer is 'yes' but we don't need a new model for the purpose. Instead, we can accommodate the dimensions of communication just mentioned by elaborating the signaling model with which we began while preserving some insights of the extended senses model. In the course

of doing so we will also be able to encompass both codification and inference.

Communication is a matter of successful signaling, which when not deceptive conveys information about the world, including, and often via, the signaler's attitude toward it. Human communication, the most complex such system known to us, fits into this broader framework at a number of levels traversing what is done involuntarily, what is allowed without being willed, what is willed, and what is both willed and done overtly.

Involuntary. Confronted with something sufficiently noxious I will be unable to help making a disgusted face. It is an empirical hypothesis, to be established only in light of advances in the evolutionary biology of facial expression, that such a response is not just evidence but also a signal of disgust. Whether or not that hypothesis is true, my disgusted face shows my disgust, which in turn may show the presence of something noxious. By contrast, the bulging vein on my forehead might show my anger, but it is implausible that it should also be a signal of that anger. That comports with the fact that a bulging vein does not appear to be designed for communication.[6]

Allowed without being willed. I experience disgust and have an impulse to show it on my face, but this impulse is not so powerful that I cannot help making that disgusted face. I might nevertheless acquiesce in that impulse and allow myself to exhibit a disgusted face. As we'll see in Chapter 5, if an involuntary facial expression is a signal, then one that I allow, in spite of the fact that I could prevent it if I chose, is most likely a signal as well.

Willed. I intentionally make a face of disgust to convey my distaste for something. According to our definition of signaling, this is a signal of disgust as well. Further, in Chapter 4 I argue that in such a case I might also show my disgust.

Willed and overt. I intentionally make a face of disgust to convey my distaste for something, but I also make clear that very intention. This overt behavior is characteristic of so-called speaker meaning, the topic of Chapter 3. In addition, we shall see that here too I might show my disgust in spite of the intervention of the will.

---

[6] We return in Chapter 5 to distinguish two different ways in which an action can be involuntary—a "can't help it" sense and a "can't do it at will" sense. For present purposes, however, the conflation will be harmless.

In all four of these cases we signal our state of mind, and unless we are being deceitful may also show that state of mind. I contend, and will spend much effort in the following chapters to establish, that analogous points apply to many other cognitive, affective, and experiential states. Just to adumbrate some of them: Perhaps my convictions on an issue are so strong that I can't help blurting them out, or instead I might acquiesce in the impulse to do so. Of course, I might also will myself to express them, and, for the more typical case, do so overtly. In all such cases, however, I signal my convictions. While experiences, like a sensation of green or a whiff of vanilla, tend not to produce characteristic facial or other involuntary behavior, I will argue that we have ways of communicating some experiences as well, in particular by showing how they feel. That is the topic of Chapter 7.

My mental states are on the whole representations of the world around me. My perception of the tiger is, at the very least, a representation of it, as are my belief that there is a tiger nearby and my fear of the tiger. What is more, in general a mental state is distinguishable into two components: The *modality* of that state—belief, fear, disgust, intention—and what that state represents, its *content*. Mental states are thus typically complex. (Some mental states may have modality with no content: A mood such as anxiety, for instance, might grip you without there being anything you are anxious about.)

When I sincerely signal my mental state, I show you one or more components of that complex. In the case of my fear of the tiger, for instance, I might show you my fear by displaying my terrified face. Or I might show you my fear of the tiger by showing you the tiger *as fearful*. I might do this by pointing to it while making a terrified face, or by drawing a picture of it emphasizing its menacing eyes and dangerous claws, or refer to it with the word 'tiger!' in a scared tone of voice. In each of these latter three cases I have shown you the tiger through my own eyes, as it were. Further, in each of these cases I have guided your reaction to some significant feature of the world without obliging you to perceive that feature. In fact, some representations that I show you might guide your reaction without the mediation of conscious deliberation on your part.

A state of mind is thus translucent in the sense that when I show it to you, you can become aware both of it and of what, if anything, it represents. By showing not just a state of affairs, but also my attitude toward it, I not only enable you to be aware of it, but also give an indication of an

apt response. So long as I can be trusted as a reliable source of information about the world, my display of tiger-fear telegraphs the propriety of fear in response to this object; my display of anger at a rival indicates the propriety of anger in response to this threat to my place in the hierarchy of our group, and so forth. My translucent display of an object thus aids you in reacting to it appropriately—by helping me to fend it off, to attack it, to eat it, to flee from it, to copulate with it, and so forth. With the articulation of the notion of showing given in the chapters to come, another thing that can be communicated and thus shown is how an emotion or experience feels. Those possessed of an appropriate level of empathy will then be able immediately, often even with no conscious deliberation, to identify with, succor, and so on, the creature who has shown them how they feel.

While deception is of course possible, it is not so hard to know of someone's emotion by perceiving one of its telltale signs—a tone of voice, a facial expression, posture, gait, or the like. Cognitive states lack characteristic perceptual components, however, and it is dubious that we can perceive them. How, then, do we show them? In general, doing so involves showing both their content and their modality. To show the content of a cognitive state we typically make use of codification. Very roughly, the semantic properties of words, together with their mode of composition, signal the content of the communicative act.[7] Any of a large variety of speech acts, however, can have a given content, so how do we signal which speech act it is? For instance, how do I signal that what I am uttering is put forth as an assertion rather than as a conjecture or a guess? Just as I might show my intention of buying your car by putting down a substantial deposit, I show my belief by sticking my neck out, by standing behind what I say. That is precisely what I do when I sincerely invoke the institution of assertion. As we will elaborate in Chapter 3, assertion and other speech acts are undertakings of various kinds of commitment. Those commitments are liabilities. This strongly suggests that assertion and related speech acts are handicaps in the sense we introduced above. I defend this claim in Chapter 3.

---

[7] This is only roughly correct because there are good reasons for denying that semantic content is sufficient on its own for determining that content. Pragmatic processes running under such labels as enrichment (Recanati 2001), explicature (Carston 2002), or implicature (Bach 1994, 2001) are often required if we are to go from the literal meaning of the words uttered to the content of the thought being expressed by the speaker.

If I succeed in making an assertion, and if I am sincere, I not only signal but also show my belief, which may in turn give you knowledge of the world. (It seems unlikely that beliefs are themselves signals.) Other devices used in different speech acts show other states of mind and other aspects of the world. A sincere promise shows my intention, awareness of which may give you knowledge of the future.

We can also signal things overtly but not explicitly, for instance when we mean more than what we literally say as in the case involving John's whereabouts considered in the last section. This form of signaling requires our addressee to infer our likely state of mind in speaking as we do. In such cases communicators depend on their addressees to infer their state of mind from the evidence they have provided.

One primary aim of this book is to establish self-expression as a matter of signaling one's thought, affect, or experience. That, in broadest outline, is how we share our point of view. Self-expression is thus coextensive with a huge range of communication in our species. It is also a pervasive feature of communication in many other species possessed of a cognitive, affective, or experiential life. (Self-expression doesn't include all human communication or signaling: for instance it doesn't include such cases as the use of pheremones, if these are indeed signals, since they aren't expressive of anything.) It follows from what I will argue that all self-expression not only signals, but also shows thought, affect, or experience. Given the terminology we have developed thus far, that suggests that many instances of self-expression are indices as we have defined that term; some, as we shall see, are also handicaps. Further, this conception of self-expression as both signaling and showing offers a picture of communication in which we lay bare what is within, but often in order to show what is going on without, as well as what to do about it. In so doing it offers an image of communicative life strikingly different from those current today. This book attempts to justify and articulate that image.

1.3. Methodological issues

I see it as among the aspirations of philosophy to raise general and fundamental questions that we are not, or at least not yet, in a position to answer with experiments, observation, or mathematical demonstration.

Philosophical scrutiny might in time sharpen those questions, and occasionally they can then be answered a priori. When that is not the case, one criterion of the success of such philosophical inquiry is that it give us a clearer view than we had before of how we might go about settling some of our questions empirically. A well-known example is the development of psychology, which began life as a branch of philosophy but which by the turn of the twentieth century had matured as an independent field of empirical investigation. Another more immediately relevant process of this kind runs from the early development of pragmatics in H. P. Grice's theory of conversation, concocted apparently in his armchair, to the very recent development of the field of experimental pragmatics.

You already know that one of the central questions of this book concerns the nature of communication, and my approach is grounded in the concepts of evolutionary biology as informed by game theory. One criterion of the success of the book is that it generate new questions about human communication from this point of view, particularly if some of those questions can be articulated in terms amenable to experimental or observational study. The thesis of certain speech acts as handicaps is an example of this sort.

It may seem surprising, therefore, that at certain points in what follows I will offer conceptual analyses or partial analyses of various concepts such as speaker meaning and self-expression. One might wonder how an a priori inquiry could have empirical aspirations. I see no conflict here, however. Just as a game theorist might give a conceptual analysis of the notion of a game and then proceed to frame questions about consumer behavior in those terms, so too we can analyze notions like meaning and expression in the hope of formulating clear and tractable questions about communication.

One dimension of this book, then, is in the tradition of philosophy-as-pre-science. Another is humanistic: I hope it is uncontroversial that understanding self-expression is, as suggested by the above quotation from Emerson, an integral part of understanding who we are. As will emerge in the following pages, humanists interested either in self-expression specifically, or, more generally, in emotion, communication, human nature, or even the arts, ignore at their peril developments in such fields as evolutionary biology, experimental psychology, and even neuroscience. At the same time, such developments don't by themselves answer the humanist's questions. Rather, we need a framework in which we can pursue clear and

precise inquiry about how, for instance, the evolution of a trait makes a difference for its communicative function; about how the behavior of this or that cluster of neurons should be thought relevant to my ability to ''feel'' someone else's pain; about how the sensory organs known as ampullae of Lorenzini found in sharks could pertain to the nature of consciousness as it is studied by philosophers. I hope to provide such a framework.

In fact, no matter your perspective you are likely to find topics discussed here that seem only marginally relevant to the ostensible subject of the book. In addition to the examples I just gave, in these pages you'll find discussions of exaptation, the muscles of the face, Neapolitan gestures, the aesthetics of music, the logic of measurement, empathy, and many other apparently far-flung topics. I ask you to bear with me as I try to explain how my subject carries us along such a winding path. I hope you come away from the experience with an appreciation of connections you may not have considered before.

Finally, many readers will have opened this book expecting me to start out by situating my position relative to that of the behemoths of the past who have discussed expression: Collingwood, Tolstoy, Dewey, Croce, Wollheim, are some names that come to mind. Such readers may be disappointed. I will make brief remarks about some of these well-known writers at various points, particularly in the footnotes. However, I do not devote a great deal of attention to them because—to the extent that I understand them—they do not seem to me to have come close enough to a correct view of our topic to merit that attention. This is in part because their concerns were primarily with aesthetics, whereas mine include aesthetics but are much broader. In some cases, such as Wollheim, these theories are also vitiated by dependence on a now-discredited psychoanalysis; in other cases, such as Croce, finding a proper interpretation of their position is a major task in itself. For these reasons, I propose a comparatively fresh start (no start is completely fresh), leaving the assessment of this work's historic antecedents for a later scholarly project.

1.4. Glimpsing ahead

In addition to inspiring what, from the point of view of the cognitive sciences, is a novel model of communicative life, our study of self-expression

offers a number of new results of which the following ten are perhaps highlights:

Beyond Saying and Showing. A philosophical tradition stemming from Gottlob Frege and Ludwig Wittgenstein, with an earlier antecedent in Rousseau, highlights a contrast between saying and showing. Aiming to counter undue reliance within philosophy on the content of what is said or thought, Frege and Wittgentsein emphasize how various aspects of language show rather than describe.[8] This book continues that tradition while offering a refinement of the notion of showing into three varieties: showing that something is so, showing that makes something perceptible, and showing how an experience, mood, or emotion feels. At the same time, I offer an account of how what we say, and the way in which we say it, also shows what's within.

Speaker meaning. Contrary to a tradition of research stemming from Paul Grice in the fifties, speaker meaning does not require intentions to produce effects on others by means of their recognition of your intention to do so. In fact, speaker meaning does not require intentions to produce effects on others at all. I argue for this in Chapter 3, and offer an alternative conception of speaker meaning that stresses overtness over audience-directed intentions. I then revisit the issue as it relates to questions of the relation of semantics to pragmatics in Chapter 6.

Speech acts. Many speech acts may be understood as forms of handicap in the sense sketched above. That implies that we may understand them first in terms of the normative status they engender for those performing them, and, in that light, in terms of the liabilities they create for the communicator as she uses them to display her state of mind. This in turn enables us to understand how speech acts characteristically express, and thereby show states of mind, and is another topic of Chapter 3.

"*Other minds*". Construing self-expression as signaling our states of thought, feeling, or experience raises the question what is involved in showing such states to others. Our answer includes the claim that in some such cases we literally make our states of mind perceptible: you can see the anger in someone's face, hear the trepidation in her voice,

---

[8] For an overview of some of their main themes see Geach 1976; a more detailed discussion of these themes in Frege may be found in Green 2001.

or feel the exuberance in her handshake. This will show that the widely debated contenders for accounting for how we know other minds, namely the Theory Theory and the Simulation Theory, do not offer exhaustive alternatives. I argue for these points in Chapter 4.

Implicature. We often mean considerably more than we say. In some cases we achieve this by means of the conventional meaning of our words, or by employing specialized grammatical devices. In other cases we achieve this by exploiting one or more conversational norms of the sort first elucidated by Paul Grice. In addition to conventional implicature and conversational implicature, however, Grice considered the possibility of non-conventional, non-conversational implicature without investigating what might fall into this category. I will argue in Chapter 4 that important cases of expressive behavior go precisely in this little-explored category. Then in Chapter 6 I argue that locutions generating so-called conventional implicata are best understood in terms of their role in expressing attitudes.

The psychology and evolutionary biology of facial expression. Experimental psychologists have taken considerable interest in facial expressions, emphasizing the human case. A striking result of research in the last few decades is the pan-cultural nature of many of these expressions. Illuminating connections have also been drawn with the neuroscience of affect. However, in recent years psychologists have been divided on the question whether facial expressions should be considered primarily as "readouts" of affective states, or instead primarily as strategically motivated actions aimed at manipulating the behavior of others. I argue in Chapter 5 for a "strategic readout" view, on which facial expressions both display affect and aim at influencing the behavior of others. In so doing I isolate shortcomings found in each of currently opposing positions.

Idiosyncrasy and conventionalization in expression. Although an important class of expressive behaviors are pan-cultural, we also express ourselves in ways that are either idiosyncratic to an individual, family, or the like, or subject to conventions local to our culture or ethnicity. An important range of cases of this sort is to be found in so-called expressive language. In all these cases we can nevertheless show what is within, albeit only in ways intelligible to those familiar with our idiosyncrasies or conventions. Explaining how this is possible is the burden of Chapter 6.

Empathy. Empathy is often construed as sharing someone else's emotion while imagining yourself in their shoes. However, I argue that while empathy does require imaginative identification, it does not require sharing an emotion. Instead, all that is required is knowing how that emotion feels and using that know-how in an act of imaginative identification. Such knowledge can be gleaned from memory, or on the other hand conveyed by works of art. One aspect of the epistemic value of art is its ability to show how an emotion feels. I develop this theme in Chapter 7.

Qualia. The qualitative dimension of our mental lives is an obsession of philosophers concerned either to debunk or defend physicalism. My concern is with the extent to which this dimension can be communicated, and more specifically the extent to which it can be expressed. I argue in Chapter 7 that we employ a variety of powerful ways of expressing, and thereby showing, what some aspect of our experience is like. This discussion exploits intuitively familiar and experimentally established congruencies among various of the sensory modalities, and congruencies between those modalities and emotions and moods.

Artistic expression. Embodying aspects of ourselves in word, deed, or artifact can also be a way of communicating, and this process of embodiment has a natural home in artistic activity. In central cases, artistic expression shows how an emotion or experience feels. In creating an expressive artifact an artist can convey knowledge of the qualitative dimension of a certain type of experience—another theme of Chapter 7.

In the next chapter I shall bulk out our subject matter of self-expression by, first, laying down twenty dicta that help to characterize it. I will at the same time characterize self-expression in a way that will help to guide and motivate its development in later chapters.

2
Expression Delineated

The word 'express' derives from the Latin root meaning 'to press out'.[1] Yet as Ogden and Richards remark, few words in philosophical or everyday parlance are as apt to bewitch the intellect.[2] In contemporary English 'express' and its cognates have a dizzying array of uses, among which are,

a. Cynthia expressed some milk during her lunch break.
b. The abnormal gene expressed itself in the cell.
c. The function can be expressed in two variables.
d. The expression 'kick the bucket' means 'to die'.
e. Corbin's express desire is not to be bothered before the meeting is over.
f. Forrest's sigh expressed her frustration.
g. The chimp's barking expressed alarm.
h. The proposed policy expresses contempt for minorities.
i. God expressed Himself in the landscape to mankind.
j. The music expresses anguish.

That a sentence like (a) could be heard in today's workplace without generating confusion suggests that we still have use for the notion of expression as pressing out. Where, likewise, would we be without espresso machines? Now, however, it is also unsurprising to hear sentences such as (c) and (d), which convey a picture of words as expressing meanings, concepts, or thoughts. Being abstract, meanings, concepts, and thoughts are not in the business of being pressed out or through anything. On the

---

[1] *Oxford English Dictionary*, Compact edn., p. 934.
[2] ''It is certainly true that preoccupation with 'expression' as the chief function of language has been disastrous. But it is not so much because of the neglect of the listener thereby induced as because of the curiously narcotic effect of the word 'expression' itself. There are certain terms in scientific discussion which seem to make any advance impossible. They stupefy and bewilder, and yet in a way satisfy, the inquiring mind, and though the despair of those who like to know what they have said, are the delight of all whose main concern with words is the avoidance of trouble'' (1923, p. 231).

other hand, words are intuitively thought to *embody* such abstract things as these to facilitate communication. That is evidently why words and phrases are also referred to as expressions, as in (d). This notion of embodiment runs through the remainder of examples (e)–(j): Corbin made her desire manifest; we can hear alarm in the chimp's bark; a theist will claim literally to see God or His image in the landscape; and so forth.

Philosophical students of language tend to focus on words, expressions, and thoughts that might represent states of affairs, but rarely embody them. There is nothing canine about the word 'dog', and nothing either cold or wet about the sentence 'Snow is falling'. We are familiar with onomatopoeia, where the sound of a word might embody the characteristic sound of the type of object or event it represents. Yet such cases seem more like curiosities than inspiration for a philosophical theory. However, recalling the discussion in Chapter 1 of the notion of a signal, let's observe that some embodiments are designed to convey information. The chimp's barking is designed, though perhaps not by him, to convey alarm, while Forrest may well have designed her sigh to convey frustration. That is why if the chimp's skin exhibits a galvanic change in response to his alarm, we wouldn't call that change an expression of alarm: even if it does embody alarm, there is no reason to think that a galvanic skin response is designed either by chimp or natural selection to convey alarm.

While some uses of 'express', such as (a) and (b) just depend on the notion of embodiment, the use that will be of most interest for us will be glossed by the idea of a *signal that embodies thought, affect, or experience*. That is found in examples (e) through to (j). We need, however, to clarify this notion of embodiment, and my first step will be to bypass some of its distracting associations by replacing it with the notion of *showing*. I will then consider the various ways in which showing can occur and the various ways in which it can be used in a signal. To that end, below I'll formulate a number of dicta that help delineate the phenomenon of self-expression. Some of these will seem too obvious to be worth mentioning whereas others will be controversial. Later chapters will then substantiate the controversial dicta and elaborate on certain of those dicta that may seem obvious. Why belabor the obvious? Some of the most everyday phenomena also give rise to fascinating explanations; a shower curtain moves toward you during your shower due to the same principle that

enables airplanes to fly, and the same principle that makes liquid climb the sides of a narrow glass is what also enables most plants to grow and geckos to climb on ceilings. So too, I hope by elucidating some of the most quotidian features of self-expression to reveal things that have been hidden in plain view.

In preparing to formulate some dicta governing the phenomenon of expression, imagine the following scenarios:

1. Jane gouges out the eyes in a photo of Judy.
2. Jane thinks about Judy, ''That jerk.''
3. Jane hugs Judy.
4. Jane sees Judy approaching and, in plain view of Judy, scowls.
5. Jane sees Judy approaching and blushes.
6. Jane says to Judy, ''I'm angry at you.''
7. Jane says to Judy, ''You jerk.''
8. Jane draws a picture consisting of vividly colored squares.
9. Jane's arm itches and she scratches it.
10. Jane predicts, ''It's going to rain.''

I shall refer to these cases in what follows.

2.1. Twenty dicta

2.1.1. A self-expression shows a thought, feeling, or experience

When in the grip of a mood, feeling, attitude, experience, or emotion there are at least three things that we might do in response. First of all, we might attempt, with greater or less success, to keep it inside. Secondly, we might act on it. Thus for instance, feeling angry at someone I might punch them; I might run at full speed from what is terrifying me; I might snatch the food I desire; I might cry out in pain. Thirdly, we might choose to publicize that mood, feeling, attitude, experience, or emotion. Whereas hiding from view is a way of acting on my embarrassment, and might even reveal it, I do not generally hide from view in order to publicize that feeling. On the other hand I might wear a sheepish look in order to make my embarrassment known. I overtly gaze at the food in order to show my desire for it. I look daggers at someone in order to show him my anger. Alternatively I can use words to show what is going on in me. (At this

point we may remain neutral on the question whether publicizing a mood, feeling, etc. is a way of acting on it.)

It is natural to describe case #4, in which Jane overtly scowls at Judy, as an instance in which Jane expresses her anger, frustration, or contempt at Judy. She could also express her feelings verbally, as in case #7. These cases suggest that one aspect of our intuitive thinking about self-expression treats it as showing, manifesting, or revealing one's thought, feeling, or experience. This phenomenon of showing, manifesting, or revealing what is within is central to the idea of embodiment mentioned above. The idea is further supported by such familiar injunctions as "Don't just tell me you love me; show me your love!" What is being called for is an expression of love not a report of it, and that expression had better embody some love, either in act or (pricey) artifact.

Calculations might show that a planet is in a certain position in relation to a star without showing that planet itself. Indeed that planet may not be visible with our best light telescopes and yet we may infer its existence. Likewise, it is possible to show that a person is in a certain affective state without showing the state in such a way as to make it perceptible. For instance, a person's failing to keep a scheduled appointment may show that she is upset, but it would not be natural to describe her absence as making her feeling perceptible. On the other hand we often suppose not just that we can determine *that* a person is in a particular state of feeling; we also often take ourselves as capable of observing those feelings with one or more of our senses. We *see* the elation spread over one person's face as they comprehend the excellent news; we *hear* the impatience in another's voice as they try to correct our misunderstanding; we *feel* the exuberance in a friend's robust handshake. Alberti admonishes the student of painting to render each face so that it shows the "movement" of its owner's soul (1954, p. 77). Charles Darwin speaks of expressive movements as revealing our thoughts and impressions more truly than do our words (Darwin 1998, p. 359). In *The Bonesetter's Daughter*, Amy Tan's narrator recounts how GaoLing approaches a fortune-teller to learn the fate of her estranged, opium-addicted husband, and GaoLing interprets the fortune-teller's pronouncement as meaning that her husband is dead. That interpretation is challenged by others, but the narrator reports GaoLing as shrugging it off: ' "Can't be," GaoLing said, but I could see a crack of doubt running down her forehead.' (2001, p. 276) We perceive all these

things—elation, impatience, exuberance, doubt—because the person is showing her feelings in such a way as to make them perceptible, and not merely showing us *that* she is experiencing those feelings.³

As we consider actions that display our feelings, we also approach the phenomenon of expression. For an illustration consider this passage from Turgenev, in which a mother gazes at her imposing but nihilistic son:

She was afraid to caress Bazarov, and he gave her no encouragement, for he did nothing to invite her caresses; and besides, Vassily Ivanovich had advised her not to "disturb" him too much…. but Arina Vlasyevna's eyes, looking steadfastly at Bazarov, expressed not devotion and tenderness alone, for sorrow was visible in them also, mingled with curiosity and fear, with a trace of humble reproachfulness. (1948, p. 152)

Let's leave aside the question whether Arina Vlasyevna's eyes could show so many different emotions at once. More pertinent, the author implicitly assumes that what is expressed in Arina Vlasyevna's eyes is also visible in them. This can only be if Arina's eyes make perceptible what they also express.

Not all self-expression makes what is within literally perceptible. An assertion expresses, and thus shows, a belief if it is sincere, but beliefs are not the sorts of things that can be perceived. Rather, a sincere assertion shows a belief by showing *that* we believe the content asserted, thereby enabling others to be aware of it; this point is developed in Chapter 3. Similarly, as we shall see in Chapter 7, in some cases our self-expression enables others to empathize with how we feel (think, experience) rather than enabling others literally to perceive it. In that case we show *how* we feel (think, experience) rather than making an aspect of ourselves perceptible; but we are still showing an aspect of ourselves.

2.1.2. A self-expression shows one's thought, feeling, or experience

Everyday usage of the notion of expression also appears to allow for the possibility of insincere expressions. On the face of it, we seem perfectly capable of expressing a regret that is not ours. However, one who harbors no regret cannot express *her* regret; this flows merely from the semantics of the genitive expression. Similarly, I am not expressing my anguish if I

---

³ The thesis that some emotions can literally be perceived is developed and defended in Chapter 4.

compose music attempting to capture the anguish of someone else. One who expresses anguish without being anguished is performing an act that is express*ive* of anguish without expressing *her* anguish. (We'll elaborate on this distinction between expressiveness and expression presently.) She may also be said to offer a putative or ostensible expression of a thought or feeling.

2.1.3. A self-expression is not a type of statement

Lessing writes, "It is a different impression which is made by the narration of any man's cries from that which is made by the cries themselves" (1949, p. 17). His point is evidently that a description of an act of self-expression lacks the visceral effect of witnessing the expression itself. This difference may be traceable to a deeper distinction between assertion and self-expression. One making a statement, as in case #10, is liable to at least the following two kinds of criticism: she may be criticized for lying, and she may be criticized for saying something incorrect. In the former case she does not believe that it is going to rain; in the latter case no rain was in the cards. By contrast, one using an unusually bright tone of voice in greeting an old school friend, as in, "Well hel*lo*! How *are* you?" may express delight in seeing her. If in fact he is not delighted to see her, he can be criticized for being insincere, phony, deceptive, mendacious, or misleading; but not for being a liar and certainly not for saying something incorrect.[4]

Self-expressions are not a form of statement (or of any of its cousins such as assertion, claim, contention, etc.). However, statements themselves have an expressive component: a statement purports to be an expression of belief, and if it is sincere it is such an expression. As a result, although all sincere statements are expressions, not all expressions are statements.

2.1.4. A self-expression is a signal

In Chapter 1 we defined a signal as any feature of an entity that conveys information (including misinformation) and that was designed for its ability

---

[4] Thus Rudolph Carnap: "The laughter does not assert the merry mood but *expresses* it. It is neither true nor false, because it does not assert anything, although it may be either genuine or deceptive" (1935, p. 28). Also, our observation that expression seems more akin to showing than stating might seem to run counter to the idea that people express their opinions. Is this not equivalent to stating their opinions? As we shall see in Section 6.3, a single act can be both an assertion of a proposition and an expression of an opinion. As such, it will follow that in a single utterance a person can both assert a proposition and show her state of mind.

to convey that information. While self-expressions show one's thought, feeling, or experience, it is not the case that all phenomena that show such things are self-expressions. As we have mentioned, a galvanic skin response might show a primate's fear, but does not express his fear. Likewise, consider the difference between blushing and weeping. A blush shows my embarrassment, but it doesn't seem natural to say that it expresses my embarrassment. Contrast this with a remark of Frederick Douglass in which he recalls the songs he and his fellow slaves would sing while not working for their masters:

The mere recurrence to those songs, even now, afflicts me; and while I am writing these lines, an expression of feeling has already found its way down my cheek. To those songs I trace my first glimmering conception of the dehumanizing character of slavery. (Douglass 1846, p. 14)

Both the tears and the blush are involuntary, and both show an emotion—grief in one case and embarrassment in the other. Yet Douglass finds it natural to think of the former as expressing his feeling. Why? I suggest this is because he intuitively sees weeping as a signal of grief, and he may well be right. By contrast, we don't intuitively think of blushing as designed to convey that one is embarrassed. It is more natural to describe weeping as an expression than it is to describe blushing as an expression because we find it more natural to construe the former as a signal than the latter.

We might of course be wrong. Advances in the evolutionary biology of facial expression might show, for instance, that blushing did evolve to signal embarrassment. In that case, consistent with the familiar phenomenon of common sense being educable, the intuitions to which I've just appealed would turn out mistaken. We would, however, still be able to understand why it *feels* more natural to describe weeping but not blushing as expressive: we would be in the grip of an (incorrect) folk theory.

2.1.5. A self-expression may be involuntary, voluntary, or both voluntary and willed

Douglass's case of the tears expressing grief is involuntary; he probably can't help that tear from forming, although he can wipe it off quickly after it does. Many things that befall us are designed to convey information but were not designed by us consciously or intentionally to do so. Again, the

terror in my voice is not something that I either will to be there or even allow to be there when I happen upon the intruder in my home. Yet that tone of voice expresses my terror because I am outfitted by natural selection to react in that way to situations of extreme danger. Behaviors that, at the time of their onset, we cannot prevent, I shall call *involuntary*.[5]

Some signals we can't help but make. Others are automatic but preventable: we don't will them to happen, although we can prevent them if we so choose. If after a brief absence I see a loved one, I am likely to smile. I don't will myself to smile, it just happens to me. Yet if I choose to do so I can suppress the smile—for instance if I see that loved one at a funeral. If I do allow myself to smile, however, I express my pleasure because here, too, we have a signal that reveals what's within. Behaviors that occur such that at the time of their onset, we can prevent them, I shall call *voluntary*.

Yet other voluntary behaviors are also ones that we will. I might be genuinely glad to see an old friend, but other things may be weighing on me so that I don't even feel an impulse to smile. (Seasoned comedians are known to watch one another's monologues completely poker-faced, reacting with no laughter or smiling at all when they find a joke funny. A good joke might get a deadpan, ''That was funny.'') I might make myself smile, however, to show my pleasure at seeing my friend, and if I do that I am still expressing my pleasure. Likewise, return to the above passage from Turgenev in which a mother's eyes express, ''a trace of humble reproachfulness''. One can only reproach a person intentionally; there are no inadvertent, accidental, or involuntary reproaches, although of course a person might *take* something I say or do as a reproach without my meaning them to. Likewise, cases #3 and #4 seem to be good examples of expression; in the first Jane expresses affection, in the second anger or annoyance, and in both cases her actions are under her control. As we will see in Chapter 4, in such cases it may also be natural to describe the agent as meaning something—that he is pleased, that her son is blameworthy, and so on.[6]

---

[5] It is of course possible intentionally to put myself in a situation in which I produce behaviors that are involuntary according to the present definition. In such cases, I can still be held morally or legally responsible for the results of those behaviors. Hence, from the fact that my behavior is involuntary in the present sense, we cannot infer that I bear no responsibility for it.

[6] The distinction between intentional and non-intentional uses of 'express' and its cognates is highlighted by Sellars 1969, who traces the distinction to ''radically different senses'' of the word

The phenomena of signals that show one's thought, affect, or experience cut across the voluntary/involuntary distinction as I have drawn it here. Since in all these cases we can show thought, affect, or experience, such signals are also cases of self-expression.

2.1.6. A self-expression can be at once spontaneous and voluntary

Although self-expression is something that we either do or allow, it is nevertheless often spontaneous. An action's being spontaneous does not mean that it springs from within and results in behavior whether we want it to or not; that is, the spontaneous is not identical with the involuntary. Being involuntary is perhaps one way in which an action can be spontaneous, but it is not the only way. Rather, an action may be spontaneous as a result of its not being premeditated or willed. That action may nevertheless be intelligible in light of a reason, that is a belief and a desire. Thus for instance when Jane hugs Judy in case #3, Jane's action might be spontaneous in that it springs from no prior deliberation. Yet we may find that Jane has a reason for doing as she does, such as that she is delighted to see her old friend after a long absence, and she could probably produce this reason with little difficulty if asked to account for her behavior. Again, an action can be spontaneous and something that we allow to happen, although it is not something that we could perform at will if we chose. As we shall see in Chapter 5, certain smiles involve the contraction of a muscle around the eye that most people cannot contract at will. In some cases we can prevent this muscle from contracting, but in those cases in which we could prevent this contraction but refrain from doing so, we may well have a reason for allowing our face into the configuration that results.

2.1.7. Although one can express only those states of ourselves that can be shown, it is an open question just what this class includes

We can only express states of ourselves, but it does not seem possible to express one's influenza or one's meningitis. Instead it seems that only states or events of a "mental" sort can be expressed. We can express fears, hopes, beliefs, and concerns. We can also express our exuberance, doubt, disdain, or dejection. The mental/non-mental distinction is, however, too crude

(p. 520). We may impute different uses to 'express' and its cognates without committing ourselves to the view that those uses have been ossified in either ambiguity or polysemy.

to delineate our topic. An itch is a sensation and might thus be considered a mental event, but it is difficult to see how in case #9 Jane could be expressing her itch. Nor would she do so if she pointedly scratched her itch, making clear to others that she intends them to see that she has an itch. It is hard to know what sense can be attached to the ideas of expressing an itch, or other sensations such as a feeling of warmth, or of dizziness. Of course, a feeling of warmth might make one irritated, and one might act crabbily as a result. However, one's crabby behavior would in that case be an expression of irritation rather than of warmth.

Some authors hold that it is conceptually impossible to express sensations such as those just mentioned, or one's perceptions such as that of a blue field, the taste of vanilla, or the smell of a lemon (Tormey 1971, ch. 1; Wright 1998). Laying down this conceptual bar goes beyond holding that it is difficult to know what it would mean to express states such as these. By analogy, we might not know what it could mean to identify consciousness with a process in the central nervous system; that does not imply that such an identification is impossible. Recalling the above dictum that self-expression requires showing one's thought, feeling, mood, or other state of oneself, the difficulty we feel in grasping what it would be to express one's taste of vanilla may instead be due to the fact that we don't know how to display this experience. For all we know, this failure may be due to a lack of skill or imagination on our part rather than to a conceptual barrier.

Perhaps visual perception can be displayed. A familiar way of understanding certain genres of painting, such as impressionism, is to see them as artists' attempts to show not a visual scene but rather their perception of it. On our gloss thus far of self-expression as showing how things are within, it follows that success in such an enterprise would be a matter of showing one's perception of a scene. Evidently, doing so requires showing how a scene looks from one's point of view.[7] This is in contrast to showing one's joy, for instance. Showing one's joy may but need not involve showing its qualitative character. One could instead show one's joy by exhibiting one of its characteristic ''parts'', such as exuberant behavior. Because a perceptual

---

[7] We'll see in Chapter 7 that if a visual perception can be displayed, this is not because that perception has itself been made perceptible (supposing that were possible). Rather, it is because the artist has enabled a third party to know what her perceptual experience is like. This is in contrast, then, to the account we'll offer of expressing one's anger in such a way as to make that anger itself literally perceptible.

state has little or no characteristic behavior (or tendencies thereto) among its components, one is left with only its qualitative character to show. One who succeeds in doing so may manage to express her perceptual state.

Some expressed states are relatively complex, such as the belief that all A's are B's. (This is expressed by a sincere assertion that all A's are B's.) Similarly, in addition to expressing relatively common emotions such as anger or fear, it is possible as well to express hope, regret, or jealousy. At the same time we can express relatively unstructured states such as moods. For instance we express dejection or anxiety, or a general malaise. Why is it that we can express these states but not itches or feelings of being warm or dizzy?

As with our hypothesis about perceptions, I suggest the answer lies in the fact that some of our mental life can be shown fairly readily, while some of its other aspects are, to say the least, rather difficult to show. What would it be like to show one's feeling of being warm, in contrast to showing that one is warm (by sweating, becoming flushed, or fanning oneself)? Again, though, it would be rash to claim that it is in principle impossible to express such a thing. One way of understanding a challenge facing artists is that of expressing aspects of inner life that do not readily lend themselves to behavioral manifestation. Anybody can express anger. It takes an artist's sensibility and technical skill to express a perception of a rose petal or a dusky seascape.

These remarks indirectly confirm our construal of self-expression as a form of manifestation of what is within. The reason is that this treatment accounts for a pattern that would otherwise seem arbitrary, namely that some mental states are readily expressed while others can only be expressed with difficulty, while with yet others we have little idea what it would be to express them.

2.1.8. A self-expression is characteristically, but not exclusively, directed toward an audience

It is natural and typical to address our expressions of thought, emotion, or experience to an audience. However, we do not always express ourselves for communicative purposes. For instance I might simply vent my rage toward the heavens, or for that matter my car, without expecting that anyone can or will hear me. Or I might utter an expletive as I accidentally slam my shin on a door frame. In these cases I am expressing my rage without directing the expression toward any audience.

*2.1.9. A self-expression may be directed toward an audience that is distinct
from the object of the state expressed*

In cases such as #3 and #4, Jane is expressing her feelings toward Judy—in
one case affection and in the other case annoyance or anger. The feelings
themselves are about a certain person, and Jane's expressions of those
feelings are also directed at the object of those feelings. However, the
object of a feeling or emotion is not necessarily the intended audience
of our expression of that emotion. One can direct one's expression of
contempt for a baseball player's botched play to a nearby fan, saying "What
a bozo!", in full knowledge that only the fan can hear the remark. The
object of the contempt is nevertheless the player, not the fan. In some
cases, indeed, our interests are best met by directing our expression of
emotion away from the object of the emotion. We do better to address
our expression of fear of the snake to those who might be in danger or
in a position to render protection, than toward the snake itself. Disgust is
often directed upon inanimate objects, which, unlike those in our social
group, don't merit the warning that an expression of disgust conveys. It is
more expedient to direct our disgust over the rotting carcass toward those
liable to get too close. It is controversial whether all emotions have objects,
and it may not be that all expressions of emotion are addressed to some
audience or other. Nevertheless we shall beg neither of these questions by
distinguishing between the *object* of the emotion and the *intended addressee*
of its expression.

2.1.10. Self-expression falls into overt and non-overt varieties

Kate Chopin's *The Awakening* describes how the main character Edna was
infatuated with an actor when she was an adolescent:

The picture of the tragedian stood enframed upon her desk. Any one may possess
the portrait of a tragedian without exciting much suspicion or comment. (This was
a sinister reflection which she cherished.) In the presence of others she expressed
admiration for his exalted gifts, as she handed the photograph around and dwelt
upon the fidelity of the likeness. When alone she sometimes picked it up and
kissed the cold glass passionately. (1972, p. 32)

Edna expresses her love for the tragedian by kissing the picture passionately,
even if she does not intend to direct that expression at the tragedian. Violet
attempts to stab her husband's dead girlfriend in Toni Morrison's *Jazz*, and

in so doing expresses rage toward that woman, among others. Similarly, in case #1, in which Jane gouges out the eyes in a photo of Judy, Jane expresses an anger whose object is Judy but the target of whose expression is Judy's photo; analogous descriptions would apply to Edna (the object is the tragedian, the target is the photo) and Violet (the object is (at least) the husband's girlfriend, the target is the girlfriend's body). In these cases, not only does the object of the emotion expressed differ from the target of the expressing act, the act's target stands in as a prop for the emotion's object. One way to understand this is in terms of make-believe: we think of Edna as making as if this picture is the tragedian, and that this kiss is landing on his lips; we think of Violet as making as if this body is the living girlfriend, and that this dagger thrust is wounding her; we think of Judy as making as if this photo is Jane, and this scissor thrust to be blinding her.[8]

These are cases of self-expression, yet they differ significantly from some others that we have considered thus far. It is doubtful that Edna, Violet, or Judy are aiming to display their emotion for or to anybody; by contrast in those other cases we may answer the question, What did the agent mean by doing what she did? In overtly scowling in case #4, Jane meant that she was angry at Judy. In calling her a jerk in case #7, Jane meant, at the very least, that Judy was a jerk. However, in gouging out the eyes in Judy's photo, there need be nothing that Jane meant. Similarly, in kissing the photo Edna does not mean that she loves the tragedian. These cases are accordingly like one possible form of case #3, in which Jane hugs Judy. In response to this hug Judy might ask, "What did you mean by that?" And Jane might reply, "I meant that I care for you." Or Jane might instead have said, "I didn't mean anything by it; I just felt like hugging you, okay?"

We will return in a moment to why one might stab a corpse or kiss or deface a photograph, but it seems clear that meaning something is not among the reasons. Accordingly, case #4 involves intentional communication, whereas case #1 does not in spite of showing, and perhaps also signaling, an emotion. In Chapters 3 and 4 we will clarify the intentions that distinguish such cases. For the present we can distinguish between *overt* and *non-overt* varieties of expression. Covert forms of expression, in which an agent

---

[8] This gloss of the function of props in terms of make-believe is not mandatory; but for brevity I will wait until Chapter 7 before considering another account of how we use them.

expresses herself with the intention of keeping that expression out of view, will be a special case of the non-overt.

2.1.11. *We can express ourselves by means of "sayings in our heart"*

Many of our thoughts, feelings, moods, and experiences go unexpressed. They might thus have objects but are not addressed to anyone. For instance Jane might feel angry at Judy without articulating that anger, to herself, Judy, or anyone else. Likewise, simply entertaining a thought to determine its logical relations to others is not thereby to express, or to purport to express, any belief or other cognitive state. Nor would a memory's occurring to one or having a tune run through one's head. On the other hand, publicizing a feeling is not necessary for expressing it. We address ourselves, and sometimes even pretend to address others, in what is sometimes referred to as 'saying in one's heart', or 'inner monologue'. In case #2 above, for instance, in which Jane thinks about Judy, "That jerk," she is articulating her feeling of anger at Judy, an anger whose object is Judy. She is also directing her expression of that anger to herself. In Ann Packer's *The Dive from Clausen's Pier*, the narrator Carrie has been visiting her fiancé in the hospital after his spinal injury from a dive off a pier. She is gradually losing interest in caring for him, and wants to move away:

Again he closed his eyes, and now tears seeped out, a single trail moving down each cheek. I set his hand down and began stroking his forearm again. I wish I could say I felt selfless then, unaware of myself. That I was thinking only of him, or that I wasn't even thinking. But I was: *This is me doing the right thing. This is me being brave and strong for Mike.* (Packer 2002, p. 102)

When Carrie thinks, 'This is me being brave and strong for Mike,' the object of her thought is a certain relation between her and Mike, but the target of her articulation of that thought is herself. She is telling herself that she is being brave and strong for Mike. In so doing she is expressing to herself a thought whose object is a relation between her and Mike. From the elucidation we have arrived at so far, this also means that Carrie is showing herself her belief that she is being brave and strong for Mike. The fact that she is doing so rather than devoting her full attention to Mike is, of course, her reason for thinking that she does not wholly believe what she is saying to herself.

In addition to self-expressions ''said in our heart'' and directed toward ourselves, some of what we say in our heart is not self-addressed. When Jane thinks to herself, ''You jerk,'' as she looks at Judy, she is expressing contempt at Judy without directing that expression at anybody, Judy, herself, or otherwise. (She is perhaps making as if to direct her contempt at Judy.) She is nevertheless expressing her contempt, and not just feeling contempt, for Judy.

2.1.12. *Self-expression is as sensitive to how an action is carried out as it is to which action is carried out*

It is not John's closing the door that expresses his anger, but rather that the closing is a slam. Mary doesn't just smile she beams, and thereby expresses her delight rather than mere happiness. Self-expression is often achieved not by the type of action performed but rather by the character of the performance. It is in this, as one might say, adverbial dimension of self-expression that it also gets a (only *a*) foothold in conventionalized forms of behavior. For once an action of a certain kind becomes conventionalized, departures from convention that nevertheless stay within the framework of the action–type in question can take on expressive dimensions. High relief often depicts heroic acts in war or conquest, with materials that project a few centimeters from the surface of the depiction. If an artist depicts a battle scene by means of figures that project significantly more, say by up to ten centimeters, she may achieve a dramatic effect expressive of anguish and terror while remaining within the confines of the genre of high relief. Another example is the crude brushwork of a painter like Dubuffet.

2.1.13. *One can express oneself in a voluntary act without intending to do so*

Return to case #1 above, in which Jane gouges out the eyes in a photo of Judy. Here Jane is expressing her rage at Judy, but is not directing that rage at Judy. Instead she is directing that rage at something that she is using as a prop, a stand–in for her former friend. Unlike an involuntary scowl, which might also express one's rage, Jane's action is within her control. Yet it is doubtful that she mauls the photo in order to express her rage; that aim would be better served by an overt, Judy-directed, scowl, or by some choice words. Why might she maul a photo, and how is her act a self–expression in spite of being voluntary but not intended as a self–expression?

A widely accepted framework for explaining and interpreting actions involves the imputation of reasons for those actions. On one common view, such a reason consists of a pair of a belief and desire. For instance, I see Mary waving from across the parking lot. We can explain this action by supposing that she wants to get my attention, and believes that by waving she may get that attention. This reason explains Mary's action by rationalizing it, showing it to be a reasonable thing to do given her beliefs and wants. Again, if Jane is furious with Joan she might physically attack her, but it's not credible that Jane thinks she'll cause Joan pain by defacing her photograph. On the other hand it does seem that Jane has a reason—*sensu* a belief and desire that, together, would make her action a reasonable one to perform—for doing what she does to the photograph.

What might that reason be? We often use props, objects that we make as if to be other things. One child may pretend to ride for the Pony Express while sitting on her rocking horse, another may pretend to walk a dog by pulling a stuffed animal on a string. The first child makes as if the toy on which she sits is a horse, the second makes as if his string and stuffed animal are a leash and dog respectively. These props make the children's imaginings more vivid than they would likely be in the absence of any physical aids. Rather than just imagining that something is so, these children can imagine those very objects to be a horse, a dog, and so on, and the physical reality of these props gives them features that are independent of the wishes of the children. For instance, the hobby horse does rock very fast, requiring the child to hang on tightly in order not to be thrown off. This will, however, make it the case that not only in actuality, but also in her imaginative game, the horse is galloping. Because vividness in experience is often due to that experience's seeming to be beyond our control, props can lend power and immediacy to our imaginative lives (Walton 1990).

Using props in the activity of make-believe is not the exclusive province of children. In the film, *Last Tango in Paris*, the character played by Marlon Brando spends long hours in the film brooding over the body of his dead wife. He often speaks to the body, and it seems clear that he is deriving comfort from doing so in spite of knowing that his wife cannot hear him. Nonetheless he finds succor by imagining that she is still there with him, and because the body before him has the same shape as the body his wife had when alive, the face is the face had by his wife, and so on,

it makes a serviceable prop. Were he to feel a surge of anger toward his wife, he might mutilate her corpse, thereby deriving some satisfaction by imagining that he is mutilating her. Similarly, in gazing at a representation of a person's face it is quite easy to imagine that we are confronting, and consequently able to affect, that person. We sometimes talk to the people we see in pictures, particularly if we have strong feelings toward them. Likewise Jane's defacing of Judy's photograph allows her to make as if she is mutilating her, and not just her photograph. This is why her action may be explained in terms of a reason. Jane defaces the photo because she desires to make as if to mutilate Judy, and she believes that by defacing the photo she can do that. If we ask why she would want to make as if to do such a thing, the answer is that making as if to do x allows one to imagine that one is doing x, and gives some of the satisfaction of actually doing x. (In the case at hand, it has the further virtue of not carrying along the legal or moral consequences of actually doing x.) [9]

Finally, Jane's action is a self-expression because, roughly, it not only shows her rage, it also was designed to do so. Not by Judy. Rather, attacking something associated with an object of rage is a widely used signal of that rage. Whether that usage is a cultural phenomenon or rather part of our biological endowment, or some of each, is a question we need not settle to understand the expressiveness of Jane's act.

[9] Hursthouse 1991 contends that Joan's behavior is a counterexample to the Humean conception of intentional action outlined above. Her reason is that Joan's behavior is not a means by which she realizes her goals, given her beliefs. Extrapolating from her remarks, this is because the only plausible candidate explanation a Humean can offer here is that Joan wants to express her rage toward Judy, and believes that by defacing her photo she can do so. Hursthouse will then point out, however, that this is not a psychologically realistic pair of attitudes to ascribe to Judy, who need not have any such self-conscious attitudes.

Smith 1998 replies to this line of objection to the Humean approach to action explanation as follows. It is a mistake, he points out, to assume that the only conative attitude that could explain her action is rage. Instead, Smith suggests, Joan acts as she does because she desires to deface Joan's photo, and believes that by puncturing the eyes of that photo with a pen she can do so. Smith acknowledges that this mode of explanation, while allowing us to preserve the letter of the Humean approach, merely prompts the question, ''And why should Joan want to deface Judy's photo?'' Accordingly a supplement to the Humean explanation is required, and Smith would offer the suggestion that one in rage is prone to seek out things she associates with the object of that rage and try to destroy them.

Smith's explanatory principles are unexceptionable, but once again we might want to ask *why* they are true, that is, why people tend to want to destroy things they associate with the objects of their rage, even when doing so will not harm that object in any way. With the notion of make-believe we may press the explanation a bit farther. The reason why we want to destroy things we associate with the object of our rage is that in so doing we may make as if to harm the person in question. Particularly in situations in which genuinely harming the person in question carries with it great costs (legal, moral, or otherwise), vicarious satisfaction of our wishes often is an appealing option.

2.1.14. Like other acts, attempts at self-expression may or may not be successful

As with many actions under our voluntary control, an agent might attempt to express her thoughts or feelings but fail to do so. It is not just that some self-expressions are more successful than others; it is also that some attempts at self-expression simply do not succeed. One obvious way in which this might happen is by her failing to make anything public; her words stick in her throat or she trips rather than pirouetting. Or what she makes public might be of the wrong kind, as her attempt at a hug accidentally turns into a cuff on the ear. More interestingly, she might produce something external, artifact or behavior, but that product might fail to express how she feels as a result of failing to show it. She feels desolate but winds up producing a picture of serenity, due to a wrong choice of color scheme. She is exuberant but the dancing is frenetic, perhaps because she is trying to do too much at one time.[10] Further, the fact that someone, who perhaps because of knowing her well, might infer from her performance that she is exuberant does not itself establish that she has *shown* her exuberance. To show, and thereby express, her exuberance, she must put it into her dancing.

What must an agent do to make public something inner in such a way that she succeeds in expressing how she thinks or feels? More precisely, how does one show one's thought, feeling, mood, and so on, in contrast to reporting it? It is relatively easy to show one's emotion when that emotion has a characteristic behavioral (including facial) signature, as is the case with disgust, fear, anger, sadness, happiness, or surprise. It is more of a challenge to express, for instance, jealousy, and painters and illustrators well know that jealousy cannot be portrayed on the face in the absence of considerable stage-setting in the picture or its title (Faigin 1990). This is to say nothing of even more subtle cases, such as the expression of a sense of lost opportunity, of power laid waste, or a feeling of vulnerability. I suggest that states such as these can nevertheless be ''put into'', that is embodied in, an artifact or bit of behavior, and thereby expressed. One aim of the present study is to explain how this is possible.

2.1.15. What is expressed, in self-expression, can be known by introspection

Many other states of ourselves can be shown besides those we have discussed thus far. I might show integrity by refraining from using information that

---

[10] The example is inspired by Wollheim 1964.

could harm a rival. However, it does not seem possible to express integrity. Why is this? The answer is that integrity is not the sort of thing that one can know by introspection: you can know by introspection of your resolve to act with integrity, or of your belief that you do act with integrity, but not of the integrity itself. By contrast, one can introspect on beliefs and other cognitive states, emotions, and experiences such as pains or sensations—and this is part of what makes it possible to express these states as well. Courage is an intermediate case. I can introspect on my determination to stand firm in the face of danger, but whether that determination amounts to courage depends on what I do in the clutch. This ambivalence seems to be reflected precisely in the fact that it is awkward, yet perhaps not absurd, to describe someone as expressing courage.

To be expressible, then, a state must be of a sort that can be known introspectively. I can know what I believe, whether I am angry or sad, or an aspect of my experience by introspection. This is not to say that every state of mind, heart, or experience that I am in is in fact known to me introspectively; there is good reason to think that such states will sometimes be unconscious. Rather, what is required for a state's expressibility is that it be the kind of state that can be known by introspection.[11]

2.1.16. *Self-expression need not take routinized paths*

Self-expression sometimes takes established routes, at other times it takes novel forms. We have conventional expressions of gratitude such as 'Thank you' as well as gestures that conventionally express contempt, approval, and disgust. Further, in recent years researchers have gathered convincing evidence that certain emotions are expressed in fairly uniform ways across all cultures. (This is a topic of Chapter 4.) Such uniformity seems not to be due to convention but rather to biological traits common across our species. On the other hand, we can also express ourselves in ways governed by no conventions or cross-cultural universals. For instance, one wishing to express her feeling of vulnerability may present someone with a delicate artifact requiring tender handling. Or you may recall this passage from

---

[11] It is misleading to speak of "the kind" of such states: any given state of myself will fall into many kinds. For instance, a single state might be a fear, a state of my central nervous system, a state caused by an imminent danger, and so forth. What is required for expressibility is that the state in question fall into *some* category such that states in that category can in general be known by introspection.

Hawthorne's *The Scarlet Letter*, when Hester Prynne first emerges from prison:

Those who had before known her, and had expected to behold her dimmed and obscured by a disastrous cloud, were astonished, and even startled, to perceive how her beauty shone out, and made a halo of the misfortune and ignominy in which she was enveloped. It may be true, that, to a sensitive observer, there was something exquisitely painful in it. Her attire, which, indeed, she had wrought for the occasion, in prison, and had modeled much after her own fancy, seemed to express the attitude of her spirit, the desperate recklessness of her mood, by its wild and picturesque peculiarity. (Hawthorne 1990, p. 53)

In what follows it will be incumbent on us to understand novel as well as routine and conventionalized forms of expression.

2.1.17. *Self-expression is distinct from expressiveness*

A bloodhound will have a sad face in spite of being perfectly content from having just devoured a large meal. Owing to an accident a person's face might be disfigured in such a way that it seems to be locked into a sneer. In having such a face he is not thereby expressing contempt. Rather than saying that his face expresses his contempt (for he may not feel any such emotion), it seems more natural to describe his face as *expressive* of contempt. This evidently means that his face has a configuration that would typically be used by one who is expressing their contempt. Similarly, a person's face might reveal something other than what it is expressive of, as in this passage from George Eliot's *The Mill on the Floss*:

It was a moment of some agitation to both, though Philip had spent many hours in preparing for it; but like all persons who have passed through life with little expectation of sympathy, he seldom lost his self-control, and shrank with the most sensitive pride from any noticeable betrayal of emotion. A little extra paleness, a little tension of the nostril when he spoke, and the voice pitched in rather a higher key, that to strangers would seem expressive of cold indifference, were all the signs Philip usually gave of an inward drama that was not without its fierceness. (Eliot 2001, Book VII, chapter VIII)

Expression and expressiveness thus should be kept firmly distinct. It also should be noted that we sometimes describe a person's face as expressive. When we do not mean this as a way of saying that the person was expressing

a particular emotion or mood, what is most commonly meant is that the person is capable of expressing a wide variety of emotions and moods in her face. Having expressive faces, as well as expressive voices and ''body language'' is a typical characteristic of good actors.

Inanimate objects can also be expressive of feelings, moods, and thoughts. A windswept cliff might look melancholy, yet we are not about to attribute feelings to cliffs. It is a large question what makes a thing expressive when it is not expressing a thought, feeling, or mood of any sentient being. Perhaps all expressiveness may ultimately be traced back to expression, and scholars have spent much effort in trying to establish just this. We shall consider this and the related question of artistic expression in Chapters 5 and 7. In the meantime, we should keep in mind that expression and expressiveness are distinct phenomena. Thus for instance in case #8 above, Jane's drawing vividly colored squares, we might describe her production as exuberant without implying that she is feeling exuberant, or that she felt exuberant when she produced the work.

2.1.18. *Corporate expression is, when successful, expressive*

Entities other than individual persons or animals are often said to express themselves. Thus for instance a corporation might express its regret for the damage it has done to an indigenous tribe's burial ground by issuing a public statement of apology directed to members of that tribe, and perhaps as well by helping to reconstruct the sacred area. For this to occur it need not be the case that any individual within the corporation expresses *their* regret for its past behavior. Further, it does not seem that the corporation itself can literally feel regret. It seems that only individual persons or animals (at most) can experience regret. As a result, ''corporate'' expressions are expressive of a thought, feeling, or attitude without expressing any individual's thought, feeling, or attitude. With this approach it will still make sense to ask whether a corporation's expression was sincere. In the case we have considered, evidence for an affirmative answer to this question will include such things as an eloquent statement of apology and a longstanding commitment to the reconstruction and preservation of the burial ground. Evidence for a negative answer will include such things as a perfunctory statement and a reconstruction project that cuts corners wherever possible.

2.1.19. Dramatic performances, when expressive, need not involve self-expression

An actor might play a character expressing grief. It does not follow that that actor is expressing grief, though this may well also be true. To heighten the drama of his performance the actor may concentrate on the recent death of a loved one and come to feel grief. That feeling may then influence his performance on stage, and he may also express that feeling. This is, however, by no means required. The actor may put in an outstanding performance without feeling any grief at all. Instead he expertly portrays a character who is himself expressing grief. Drawing on the distinction we made above, part of what makes the portrayal successful will often be that the actor uses techniques that are expressive of grief, and these will include how he carries himself, his tone of voice, and his facial configurations among other things. All these techniques may be used to build an expressive performance without the actor expressing any emotion.

2.1.20. It is an empirical question where self-expression is found in the animal kingdom, and of its ontogenesis in any given species

We have already urged that self-expression is an integral part of human life, and this theme will be developed in later chapters. To many it will also seem undeniable that non-human animals express themselves. However, as we develop our account of self-expression, both the overt and non-overt varieties, we shall see that it no longer goes without saying that a dog's wagging tail, for instance, or a silverback gorilla's piloerection, are self-expressions. The complexity of the issue is analogous to the difficulty of establishing that non-human primates have a language in anything like the sense of human language. Current debates pertaining to "animal language", I hope to show, are polarized by a shortage of choices concerning what could count as meaning, and, indirectly, expression, and I would like to contribute to these debates not, primarily, by taking a stand on one side or another but by articulating more nuanced theoretical options.

2.2. A characterization of self-expression

We are now in a position to collect together the considerations adduced thus far into a gloss of the notion of self-expression that we both elucidate and treat as a working hypothesis in the following chapters.

Characterization of Self-Expression
Where A is an agent and B a cognitive, affective, or experiential state of
a sort to which A can have introspective access, A expresses her B if and
only if A is in state B, and some action or behavior of A's both shows
and signals her B.

According to this characterization, all self-expression involves showing
one's emotional, cognitive, or experiential state, but not all such acts are
done with the intention of showing that state. For instance, I may not
laugh for the sake of showing my pleasure, but if I do laugh as a result
of feeling pleasure, and if in fact laughter evolved as a device for showing
one's pleasure, then my laughter expresses pleasure. Again, in cases in
which the target of one's expression stands in as a prop for the object of
one's expressed state (such as when Edna kisses the photo of the tragedian),
one need not act with the intention of showing how one feels. Instead one
might simply intend to make as if to act on that feeling.

Also, strictly speaking, according to our account thus far, A need not
have been in state B before her act of self-expression; it is consistent with
the account given thus far that what is expressed is somehow constituted
by the expressive act. Further, her act of self-expression needn't have been
the subject of planning or deliberation before that act commenced; it might
have been entirely spontaneous. In addition, the act in question may be an
allowing of an urge, inclination, or impulse to take its course, as opposed
to her willing herself to behave in a certain way. Further, to accommodate
the above observation that self-expression is often achieved by virtue of
the manner of the performance of an action, we note that our notion of
an action may be used broadly enough so that, for instance, closing the
door violently, or shaking hands exuberantly are both actions. As well, an
actress who dredges up a genuine feeling of despair in playing the role
of a character expressing despair may well express her own despair in the
process. She will do so if in playing her role she makes as if to act on her
own emotion—that is, an emotion of which she is in fact possessed—and
if she does this she will both express an emotion and portray a character
doing so.

Our characterization of self-expression is also meant to accommodate the
possibility of a feigned expression of emotion serving as an expression of
that very emotion. In another passage in A. Packer's *The Dive from Clausen's*

Pier (Knopf, 2002), the narrator has just described her housemates Simon's and Greg's effusions about the opulence of a party from which they've just returned. Her boyfriend directs biting sarcasm toward them. In response,

"Well," Simon said. "On that note, I think I'll go watch TV." He made a face, a sort of ironically freaked-out face that was meant to say that he actually was freaked out, and then he left the kitchen. (p. 169)

Here Simon is expressing his feeling of shock by putting on a dramatically shocked-looking face. It seems clear, furthermore, that he is making the face to manifest his shocked feeling. Simon is not making as if to act on his shocked feeling, nor do we need a new category of showing one's thought or feeling by making as if to do so. Rather one way of showing one's feeling is to make as if to do so, while accentuating or caricaturing its physical manifestation in such a way as to make clear that one's action is voluntary.

In this chapter we have offered a preliminary delineation of the notion of expression. A primary aim of the next two chapters will be to relate that notion to the broader phenomenon of meaning. We shall see that self-expression deserves as central a place in the philosophy of language and mind as do such well-entrenched phenomena as representation, reference, assertion, and implicature. Another aim of those chapters, mandated by our gloss of self-expression as embodying what is within, will be the development of the notion of showing one's state of thought or feeling. This will be equally relevant to communicative and non-communicative varieties of self-expression. For convenience the twenty dicta elucidated thus far are gathered below.

Twenty dicta about self-expression

2.1.1. A self-expression shows a thought, feeling, or experience.

2.1.2. A self-expression shows one's thought, feeling, or experience.

2.1.3. A self-expression is not a type of statement.

2.1.4. A self-expression is a signal.

2.1.5. A self-expression may be involuntary, voluntary, or both voluntary and willed.

2.1.6. A self-expression can be at once spontaneous and voluntary.

2.1.7. Although one can express only those states of ourselves that can be shown, it is an open question just what this class includes.

2.1.8. A self-expression is characteristically, but not exclusively, directed toward an audience.

2.1.9. A self-expression may be directed toward an audience that is distinct from the object of the state expressed.

2.1.10. Self-expression falls into overt and non-overt varieties.

2.1.11. We can express ourselves by means of "sayings in our heart".

2.1.12. Self-expression is as sensitive to how an action is carried out as it is to which action is carried out.

2.1.13. One can express oneself in a voluntary act without intending to do so.

2.1.14. Like other acts, attempts at self-expression may or may not be successful.

2.1.15. What is expressed, in self-expression, can be known by introspection.

2.1.16. Self-expression need not take routinized paths.

2.1.17. Self-expression is distinct from expressiveness.

2.1.18. Corporate expression is, when successful, expressive.

2.1.19. Dramatic performances, when expressive, need not involve self-expression.

2.1.20. It is an empirical question where self-expression is found in the animal kingdom, and of its ontogenesis in any given species.

3
Showing and Meaning

I have begun to characterize self-expression by situating it within the larger framework of signaling, which can occur in the absence of any intentions to send a signal or consciousness of doing so. Self-expression does require what is expressed to be a cognitive, affective, or experiential state, but it is not true that all such states are expressed intentionally. Nevertheless, one intuitively familiar form of self-expression occurs in cases that philosophers refer to as 'speaker meaning' (née 'non-natural meaning'), and such cases do involve intention. No current theory of speaker meaning is, however, accurate, so in this chapter I'll offer a new one. In particular, I lay the groundwork for an account of the relation between self-expression and speaker meaning, paying particular attention to the extent to which and the ways in which intentions are involved. I first (3.1) distinguish among three species of showing. Then (3.2) on that basis I further develop the signaling model of communication outlined in Chapter 1, explaining three ways in which signals of the sort that concern us in this book can show what they signal. I next (3.3) argue, contrary to widely held opinion, that speaker meaning does not require intentions to produce effects on an audience, much less intentions to produce effects by means of recognition of those intentions. We shall see instead that it's necessary and sufficient for speaker meaning that one overtly show something, or overtly show that something is so, or overtly show one's commitment to a content in a certain way. In 3.4 I develop this notion of overtness on which that characterization of speaker meaning depends. In 3.5 I defend a view of many familiar speech acts as handicaps in the technical sense of that term. This will in turn motivate a generalization of the notion of speaker meaning achieved up to that point. I also consider, in 3.6, some alternative conceptions of speaker meaning, and show why they are inadequate. These ideas will lay the groundwork for Chapter 4, in which certain forms of self-expression emerge as species of speaker meaning.

3.1. Three ways of showing

Our inquiry will make heavy use of the notion of showing, which comes in at least three forms. First of all, I might show my courage by acting bravely. My brave behavior is good evidence of my courage. Or I might, by means of extensive calculations, show that there is a black hole in the center of the Milky Way. In these cases I don't make what I show perceptible; I certainly couldn't make the black hole perceptible, and it is not clear what it could mean to perceive courage. Rather, in these cases I provide compelling, though not necessarily conclusive, evidence for a conclusion that could be grasped even by someone with no capacity for vision or other sensation. A grammatical tag for this category is *showing-that*. Because my brave behavior (calculations, etc.) is good evidence of my courage (the existence of the black hole, etc.), an appropriately situated thinker aware of that evidence is in a position to know of my courage (of the black hole, etc.). Showing-that thus enables propositional knowledge.

Just as a horse can be led to water without being made to imbibe, so too enabling propositional knowledge is not the same thing as guaranteeing that all interested parties will come to know what has been shown. I might prove a theorem for my class, showing them that, say, $a^2 + b^2 = c^2$, where a, b, and c are three sides of a right triangle, without everyone's cottoning on. A week later one of my students might finally get it, remarking, ''Oh! Last week he showed us that $a^2 + b^2 = c^2$; I just didn't follow the reasoning at the time.'' Making knowledge available doesn't guarantee its transmission. On the other hand, my running through the proof doesn't make knowledge available to everyone who watches. If the aforementioned horse happens to have witnessed the proof, I won't have shown him the truth of Pythagoras' theorem; the same goes for my newborn daughter. In order to show you that something is so, you need, at the very least, the conceptual resources required to believe the proposition in question. I can therefore show my student that $a^2 + b^2 = c^2$ without, in so doing, showing my horse or daughter that this is so even if they're watching.

A second form of showing makes a thing perceptible. I show my bruise, and thereby enable others to see that bruise. Although it is most natural to speak of showing in visual terms, showing is not limited to vision: one can show someone a rough texture (you'd need to feel the texture)

or a coyote's howl (you'd have to hear it). Just as what I show you in the propositional-knowledge enabling case requires appropriate conceptual resources on your part, what I show you in this perceptual case depends on your perceptual capacities and your position in the environment. If you had electroreception like a hammerhead shark, I could show you the electrical activity in the body of a fish hiding under the sand. In that case you'd not only perceive the fish, you'd "electroreceive" it. Likewise, even if there are mice in the field, I don't show you them from an airplane passing two hundred yards above the field. On the other hand if you had the visual acuity of a hawk, I might well do so. Let us put this perceptual-knowledge enabling form of showing under the rubric of *showing-α*, where 'α' is a singular term referring to a perceptible object or affair.

Finally, I might also *show how* something looks, feels, sounds, and so on. Apply friction to a scratch-and-sniff picture of a skunk. You won't thereby smell any skunk, but if your nose is functioning properly, you will learn how skunks smell. By accurately painting Mary's profile you will show how Mary looks in profile—what she looks like from that angle—thereby enabling me to know how Mary's profile looks. If I then acquire the knowledge that has been made available, I can later manifest it by reliably discriminating the Mary-like profiles from the rest. Similarly, the trepidation in my voice might enable you to know how my anxiety feels if you are sufficiently empathetic. If you are sufficiently empathetic, then hearing my voice may enable you to imagine feeling my trepidation. If you can do that, then you know how I feel. Showing-how can provide qualitative knowledge for those with appropriate sensory capacities. It can also enable empathy for those with the capacity for empathy. The above three forms of showing—showing-that, showing-α, and showing-how—enable propositional knowledge, perceptual knowledge, and either experiential knowledge or empathy, respectively.

Very often a single experience will involve all three forms of showing. My location and sensory capacities will enable me to see the cloud; doing so might show me that a storm is brewing, as well as what a cumulonimbus looks like. On the other hand, such phenomena as blindsight suggest that an event might show α without showing me how α looks; analogous things may be possible for other senses. We have, further, already mentioned that one can attain propositional knowledge of a state of affairs without perceiving it. For these reasons, clarity is best served by keeping our three

forms of showing distinct even though in typical cases they arrive bundled together.

One thread that unites the above three forms of showing is knowledge: Evidence enables those who are shown the things mentioned above, and who are in the right circumstances (being empathetic, being in the right perceptual location, possessed of the right conceptual resources or background knowledge, etc.) to know some fact, some object of perception, or how some emotion, mood, or experience feels. Showing is thus a stronger relation than indication, in two ways. First, showing, unlike indication, is a ''success'' notion: One can only show facts (showing that), or real things (showing α), or how something appears or feels (showing how), whereas one can indicate that something is so when it is not, or indicate an object that is not, or indicate how something appears or feels that does not appear or feel that way. Another thread unifying the three types of showing is that each of the forms of knowledge it enables can be made available by design. That is to say that each of the three forms of showing can be the content of a signal. Just as the height of the tiger's scratch marks shows its size, so too the intensity of my cry, when I'm neither faking nor exaggerating, shows the extent of my terror. A sincere speech act of mine might show my belief; the cry might also show how my terror feels. I develop these points in the next section.

3.2. Showing what's within, part i

Recalling terminology of Chapter 1, a *signal* is any feature of an entity that conveys information (including misinformation) and that was designed for its ability to convey that information. That information might pertain to how things were, are, will be, or ought to be. The design in question might be due to the work of intelligent agents, or be the product of evolution by either artificial or natural selection. When a signal succeeds in conveying the information for which it was designed, communication takes place. This is the *signaling model of communication*. I argued in Chapter 1 that this signaling model can accommodate the use of codes as suggested by the code model, and inference as suggested by the inferential model, as special cases of a more general pattern. In addition, I have been developing a position

that both articulates and clarifies what in Chapter 1 we called the extended senses model of communication.

For any organism and characteristic ecology, there will be a cost involved in sending a signal. The warningly colored tree frog expends some resources in producing bright colors. However, those resources appear to be negligible in relation to the resources required to produce any other skin color. For that reason, the tree frog only pays what is called an *efficacy cost* in producing its warning signal (Maynard Smith and Harper 2004, p. 15). Mimics, who wear bright skin color without being noxious, pay the same efficacy cost to produce the same signal. It should be plausible, however, that if too many mimics are found in a given environment, predators will start to ignore the warning signal and start eating all brightly colored frogs whether they are noxious or not. The result is that, when mimicry is possible, and organisms can gain from mimicry without incurring any substantial cost, signals are liable to lose their credibility.

This conclusion might make one wonder if any signaling system will be unstable, always liable to overthrow by mutations or other tricksters that produce deceptive signals. The answer is that some signaling systems seem to have developed that are resilient against such threats. We know of three conditions under which this stability emerges, where the third is a special case of the second. One occurs when the organisms in question share the same preference ordering. We have evidence that the African honeyguide bird (*Indicator indicator*) has been helping humans to find beehives for at least 20,000 years. These birds gain the attention of humans with a distinctive chirp, and the humans follow them to the beehive. The humans then extract the honey, and the honeyguides eat the larvae and wax from the nest (Isack and Reyer 1989). Neither party to this interaction would have an incentive to signal ''dishonestly'', and that's good reason to believe that the honeyguides' signals are and will remain reliable.

Another source of signaling stability is found when the signals in question are indices—signals that can only be faked with great difficulty due to limitations on the organism. We have already seen this in the example of the tiger's scratch-marks. Another case is funnel-web spiders, *Agelenopsis aperta*, who find themselves in contests over webs. Two spiders will vibrate on a disputed web. Reichert (1978, 1984) found that if two contesting spiders differ in weight by 10 per cent or more, the lighter spider retreats 90 per cent of the time rather than fighting. A losing spider can, in addition, be

made into a winner by placing a weight on its back. This strongly suggests that vibrating on a web is a spider's signal of its size. Further, in the absence of scientists placing weights on their backs, funnel-web spiders can't fake these signals. The "vibrating game" thus exploits an index of spider size. In so doing, its players not only signal their size; they also show it. It is, I suggest, precisely by showing their size that spiders have happened upon a stable signaling system.

A special case of an index is a handicap—which is a signal that is more costly than is required just for sending a signal with that content. Intuitively, when a signal is a handicap, only those who are "honest" can afford to pay the price required to send them. As we observed in Chapter 1, the ostentatious feathers of the male peacock don't just signal his viability—they show it because in spite of carrying such a handicap, he survives: the fact that he has survived up to now with all that baggage shows his viability. Another example of a handicap is found in the male stalk-eyed fly, *Cyrtodiopsis dalmanni*. In choosing mates, females prefer males with long eye-stalks (Wilkinson and Reillo 1994; David, *et al* 1998). This is in spite of the fact that these appendages make the male slower and easier to spot for predators, and are costly to produce. Observing that these appendages provide evidence of the male's viability—since he'd have to be unusually fit to survive with them—Maynard Smith and Harper (2004, pp. 33–4) hypothesize that the eye-stalks are also signals of the male's viability. A signal, such as the peacock's tail or the stalk-eyed fly's stalks, whose cost goes beyond its efficacy cost carries a *strategic cost* (Maynard Smith and Harper 2004, p. 14). My suggestion now is that when a signal carries a sufficiently high strategic cost, it not only signals some property of the organism; it also shows that property. Further, it is its ability to show this property that vouchsafes the stability of this signaling strategy.

Signaling systems, then, can become reliable by virtue of the confluence of interests of the communicators, or by exploiting indices, of which a special case is a handicap. Indices, I have suggested, gain their power by their ability to show what they index, and handicaps show what they do in a peculiar way. (Figure 3.1 depicts relations among such concepts as cue, signal, showing, index, and handicap.) Because of the intimate connection between reliable signaling and knowledge, we are now in a position to see that reliably signaling a state of affairs enables others who are appropriately
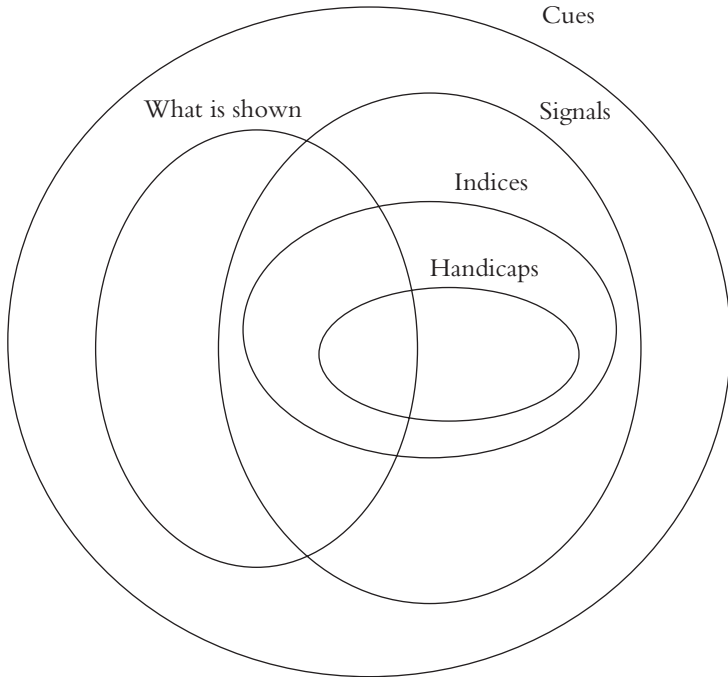
Figure 3.1: Relations among cues, signals, showing, indices, and handicaps

situated to know of that state of affairs or an aspect of it—either by knowing that it obtains, by knowing how it feels, or simply by perceiving it.

We've now arrived at a picture of reliable signaling as, inter alia, making knowledge available to others. Recalling our characterization of self-expression from Chapter 2 as signaling and showing some aspect of one's point of view, this implies that in expressing ourselves we make knowledge of our point of view available to others. To make that picture plausible we will need to explain more fully what it is to show one's cognitive, affective, or experiential state, and that will be an aim of the coming chapters. To make the picture informative, we will need to explain how it relates to other notions that are common fare for students of language, communication, emotion, and experience. In particular, the picture we've arrived at appears far removed from the kind of signaling that for many readers will seem ubiquitous in our own species, namely, the kind that exploits ''speaker meaning''. A half-century-old tradition in the philosophy of language, for instance, has it that the kind of ''meaning'' carried by

a tiger's scratch marks, a fly's eye stalks, or a peacock's tail feathers is profoundly distinct from that carried by a person's intentional honking of a car horn or, for that matter, her utterance of words. It may, further, seem that in describing speaking as a way of signaling, we only scrape the surface of the subtleties with which linguists and philosophers have wrestled for many decades. I'll argue, however, that this signaling picture is the beginning of wisdom about many aspects of verbal communication. Establishing this will span a few chapters, and our first step in that project will be a reconsideration of the notion of speaker meaning just mentioned.

3.3. Grice's ladder

Philosophers have spent great effort developing accounts of various kinds of communicative act. In this section I discuss H.P. Grice's well known approach. Grice couches this discussion in terms of the notion of meaning rather than that of communication, since the latter requires that information be not only sent but also received by its recipient, whereas the former does not and is thereby simpler. (A person might mean something without being understood, or even heard; in such a case we have an instance of meaning but not of communication.) The notion of meaning nevertheless contains many of the core elements needed for our study. As we proceed, we will find ourselves taking issue with the bulk of Grice's major contentions about the nature of meaning. It may for this reason be surprising that we are discussing his work at all. The reason is that although Grice's views about meaning are flawed in many ways, they are clues to a more adequate account that most effectively emerges once we see how he and other predecessors have erred.

   In his influential 1957 article, Grice distinguished between two senses of 'mean'. One sense is exemplified by remarks such as 'Those clouds mean rain,' and 'Those spots mean measles.' The notion of meaning in play in such cases Grice dubs 'natural meaning'. Grice suggests that we may distinguish this sense of 'mean' from another sense of the word more relevant to communication, exemplified in such utterances as:

   in saying, "You make a better door than a window", George meant that you should move,

and

    in gesticulating that way, Salvatore means that there's quicksand over there,

and

    in saying, "Look!", Alfonso means the man with the red hat,

and

    in pointing in that direction, Mary means the one on the top shelf.

Grice used the term 'non-natural meaning' for this sense of 'mean', and in more recent literature this jargon has been replaced with the term 'speaker meaning'. (This terminology is misleading because according to philosophers' usage, an act can be one of speaker meaning with no sounds uttered or even any inscriptions made. For instance two hunters with no common language might communicate with pantomime, so that when one acts out the path of attack he means, in the sense of speaker meaning, that the other is to approach the mammoth from behind. In spite of the misleading nature of the jargon of speaker meaning I shall retain it rather than introduce new nomenclature.) Although he does not discuss the point, Grice's idea of speaker meaning can be expressed in either of two forms, exemplified above. One is that in which a speaker means *that* something is the case. Here what is meant is a proposition, and that proposition may be being put forth with the force of assertion or one of its cousins such as conjecture, prediction, or supposition. Call this *propositional speaker meaning*. (In the course of this chapter, this notion will fission into what we shall call 'factual speaker meaning' and 'illocutionary speaker meaning'.) Another form that speaker meaning may take has to do with an agent having something in mind. When Alfonso says 'Look!', pointing at an eagle passing overhead, he means a particular object without, as such, saying anything about the object in question. For this reason he is not asserting, conjecturing, or performing any other illocutionary act involving a proposition. Call this sort of case *objectual speaker meaning*. Both propositional and objectual speaker meaning are familiar from everyday communicative life.[1]

---

[1] Speaker meaning is usually construed as taking only propositional objects. Even when the force of an utterance is not within the assertion family (which includes assertion, conjecture, supposition, prediction, and so forth), but is rather imperatival or interrogative, the content of that utterance is normally taken to be propositional. We nevertheless have a robust intuitive feel for objectual speaker meaning, and we can take objectual speaker meaning seriously without turning the clock back to a

After distinguishing between natural and (what we shall hereafter call) speaker meaning, Grice asks what features would enable a situation to be a case of speaker meaning. Grice first considers the suggestion that it is sufficient for speaker meaning that a person do something that influences the beliefs of an observer. This is clearly inadequate, however, since in putting on a coat I might, unbeknownst to me, lead an observer to conclude that I am going for a walk. Yet in such a case it is not plausible that I mean that I am going for a walk in the sense germane to speaker meaning. Nor do I mean any object, such as myself, my walking, or the prospect of my walking. (I might of course *mean to* go for a walk, but intending to do something is not, by itself, enough for speaker meaning either.) Thus performing an action that influences someone's beliefs is not a sufficient condition for speaker meaning, be it objectual or propositional.

Might performing an action with an intention, successful or not, of influencing someone's beliefs be sufficient for speaker meaning? It is not. I leave Smith's handkerchief at the crime scene to make the police think that Smith is the culprit. Here I leave the handkerchief where I do for the sake of influencing the beliefs of the police. However, whether or not I am successful in getting the authorities to think that Smith is the culprit, in this case it is not plausible that I mean that Smith is the culprit. Similarly, it is not plausible that in leaving the handkerchief where I do, I mean Smith, or Smith's guilt, or any other object. Accordingly, performing an action with an intention of influencing someone's beliefs is not sufficient for either objectual or propositional speaker meaning.

Perhaps what is missing in the handkerchief example is the element of overtness. This suggests another criterion, namely that of performing an action with the, or an, intention of influencing someone's beliefs, while

crude semantic theory typical of eighteenth- and nineteenth-century empiricism, according to which the paradigmatic unit of meaning is the noun.

One striking example of the use of objectual speaker meaning in contemporary literature is in Chuck Palahniuk's *Choke* (2001). Here the narrator remarks how his friend Denny has found a newspaper ad apparently placed by his parents, ''Next to me, Denny reads: 'Free to good home, twenty-three-year-old-male, recovering self-abuser, limited income and social skills, house trained.' Then he reads a phone number. It's his phone number. 'It's my folks, dude, that's their phone number,' Denny says. 'It's like they're hinting.' He found this left on his bed last night. Denny says, 'They mean me.''' (p. 122)

intending that this very intention be recognized. Grice contends that even here we do not have enough for speaker meaning. He offers the example of Herod, who presents Salome with St John's severed head on a charger, intending that she discern that St John is dead and intending that this very intention of his be recognized. Grice observes that in so doing Herod is not *telling* Salome anything, but is instead deliberately and openly letting her know something. Grice concludes that Herod's action is not a case of speaker meaning either. The problem is not that Herod is not using words; we have already considered hunters who mean things wordlessly. The problem seems to be that to infer what Herod intends her to, Salome does not have to take his word for anything. She can see the severed head for herself if she can bring herself to look. By contrast, in its central uses, telling requires a speaker to intend to convey information (or alleged information) in a way that relies crucially upon taking her at her word. Grice appears to assume that at least for the case in which what is meant is a proposition (rather than a question or an imperative), speaker meaning requires a telling in this central sense. What is more, this last example is a case of performing an action with an intention of influencing someone's beliefs, even while intending that this very intention be recognized; yet it is not a case of telling. Grice infers that it is not a case of speaker meaning either.

Grice holds that for speaker meaning to occur, not only must one (a) intend to produce an effect on an audience, and (b) intend that this very intention be recognized by that audience, but *also* (c) one must intend this effect on the audience be produced at least in part by their recognition of the speaker's intention. I shall return in a moment to his reasons for introducing this third element. Before doing so, however, let's pause over these cases in which one overtly manifests or displays an object, situation or state of affairs. In a similar spirit to that of Grice, Clark (1996, p. 129) remarks that it would be unnatural to describe a person, Ernest, who throws open a window to show Matthew the rain outside, as meaning that it is raining outside. Let us assume that as in the Herod–Salome case, Ernest's action is overt. Thus rather than, say, causing the window to be open without enabling Matthew to discern that Ernest is the one who opened it, Ernest opens the window in full view of Matthew, where it is also clear that Ernest can see Matthew, that Matthew is aware that Ernest can see him, and so on. Even with the example thus clarified it may not seem

entirely colloquial to describe Ernest as meaning that it is raining outside (or some objectual analogue). This, however, could be due to the fact that so describing the situation seems roundabout in contrast simply to describing Ernest as showing the rain outside. It would be a bit like describing a person as trying to swim across a river when he in fact is succeeding in doing so; it is true that he is trying to swim across, but this understates the situation and so can be misleading. Likewise, the awkwardness of describing Ernest as meaning that it is raining outside may be due to the fact that for many conversational purposes we could more informatively describe him as showing that it is.

To circumvent this distracting feature, then, consider a case in which what is meant is not obvious in what is shown. Adverting to what is meant will then not be apt to mislead. Accordingly, consider a variation of Clark's case in which Ernest overtly opens the window to reveal threatening skies. He shows the looming weather, and yet Matthew, not appreciating the impending storm, might be puzzled, asking, "What's your point?" Alice might intercede, saying, "He means there's a tornado over there. Can't you see the funnel cloud?"

In this last case we have an act that shows an impending tornado that can at the same time be described as the agent's either meaning that tornado (objectual speaker meaning), or that the tornado is impending (propositional speaker meaning). More precisely, a single act can be both a case of speaker-meaning that P (or speaker-meaning α, for the objectual case), and a case of showing that P (or showing α). Hence showing does not preclude speaker meaning. A variation of the Herod–Salome case makes an analogous point. Suppose that Salome is seriously nearsighted and does not trust her vision to determine whether what she sees is a severed head, or John wearing a charger-shaped necklace with the rest of his body occluded. She might accordingly not only wonder, "What is Herod showing me—a dead John or a partly occluded John?" She might equally wonder, "What does he mean? That John is dead or merely that this shackle is too tight?" The window example did not employ a telling that crucially involved taking someone at their word, for Ernest showed Matthew the impending storm even if Matthew did not immediately grasp the significance of what he had seen. Similarly, this revised Herod–Salome case does not involve a telling in this sense, for here again Salome does not need to take Herod's word for anything to determine that John is indeed no longer. She need

only look closer. After doing so, she has her answer: "Ah. He means John is dead."²

Other cases point to a similar conclusion. Imagine that A asks B for a game of squash, to which B replies by showing an extensively bandaged leg that clearly incapacitates him for sports (Schiffer 1972, p. 56). Here it seems natural to say that B means that he cannot play. Yet here B presents the leg with the, or an, intention of influencing A's beliefs, while intending that this intention be recognized. Here it does not seem that he need have any more complex intention than this, such as an intention to produce a belief at least in part by recognition of this intention. Again imagine a mathematics teacher who, at the end of a demonstration of the Pythagorean Theorem, utters, "And so, $a^2 + b^2 = c^2$." The teacher has shown that $a^2 + b^2 = c^2$ with the intention of influencing her audience's beliefs while intending that this intention be recognized. At the same time it is colloquial to describe her as meaning that $a^2 + b^2 = c^2$.

---

² Grice might also be being misled by the assumption that little of communicative interest lies between the criteria he proposes for natural meaning and those he proposes for speaker meaning. In articulating natural meaning, Grice gives five conditions that must be satisfied:

1. For natural meaning, 'x meant that p' and 'x means that p' entail p. (One cannot consistently say, "Those spots mean measles, but he hasn't got measles.")
2. For natural meaning, 'x means that p' does not entail that something was meant by x.
3. For natural meaning, 'x means that p' does not entail that somebody meant something by x.
4. For natural meaning, 'x means that p' cannot be reformulated as 'x means that "p"'.
5. For natural meaning, 'x means that p' can be reformulated as 'the fact that x occurred means that p'.

He goes on to argue that in cases of speaker meaning all five of these conditions are violated. However, Grice gives us no reason to believe that these five conditions stand or fall together. Instead he simply asserts, "I do not want to maintain that *all* our uses of 'mean' fall easily, obviously and tidily into one of the two groups I have distinguished; but I think that in most cases we should be at least fairly strongly inclined to assimilate a use of 'mean' to one group rather than to the other" (p. 215).

Return now to the case of the bandaged leg. Letting x be 'Jones presents his heavily bandaged leg' and P be 'Jones is unable to play squash', we may see that while condition 1 above is met, conditions 2 and 3 are not. 'Jones's presenting his heavily bandaged leg means that he cannot play' *does* entail that something was meant by Jones's presenting his heavily bandaged leg. 'Jones's presenting his heavily bandaged leg means that he cannot play' *does* entail that somebody (namely Jones) meant something by Jones's presenting his heavily bandaged leg.

Here then is a case showing that Grice's five conditions do not stand or fall together, but the case is by no means recherché. I conjecture that Grice failed to see that cases satisfying some but not all of the five conditions are relatively common; with that background, in seeing that the Herod case and its ilk do not violate all five conditions as do his showcase examples of speaker meaning, he inferred that they must merely be cases of natural meaning. We are now in a position to see that this was a false step. (See Denkel 1992, and Wharton 2002, for further discussion of intermediate cases such as those we have considered.)

Many theorists take Grice's main result in 'Meaning' to be a treatment of speaker meaning that involves a certain complex communicative intention. After (in our view incorrectly) rejecting the Herod case as exemplifying speaker meaning, Grice asks us to consider the difference between Z showing Mr. X a photo of Mrs. X and Mr. Y in a love embrace, and Z presenting Mr. X with a drawing depicting the same situation. As one might expect, Grice denies that Z's presenting the photo is a case of speaker meaning, and this position need not detain us. More important is his analysis of the drawing case. For Grice, if the presenting of the drawing is to be a case of speaker meaning, then Z must be doing so not just with the intention of making Mr. X believe there is an affair, but must also intend to make him believe this by means (at least in part) of Mr. X's recognition of this very intention. That is, Grice holds that Z must intend that Mr. X reason, on some level: ''Z is trying to tell me something, namely that things are as his picture depicts them. So, since Z is evidently trustworthy, I should conclude that that is how things are.'' The intention to produce a belief or other attitude by means (at least in part) of recognition of this very intention has come to be called a *reflexive communicative intention*. Figure 3.2 summarizes the progression of analyses and counterexamples alleged to motivate the requirement of reflexive communicative intentions.[3]

It is important, perhaps crucial, to communication that we take the words of others, often though of course not always, at face value. As we will see more clearly in Section 3.4, the possibility of such speech acts as telling and promising would be hamstrung if we did not do so. On the other hand, we should resist the temptation to overstate this point, by for instance taking it to be essential to speaker meaning, either propositional or objectual. Cases of speaker meaning do not always involve the intention that others rely upon our word; this has already been shown by the examples of the bandaged leg and the proof of the geometrical theorem. The same goes for a person who takes an unpopular stand, espousing her beliefs without any intention that her audience will come to agree with her. Indeed, the Pythagorean theorem example shows that in some cases we decidedly do not want others to rely on our word. The geometry teacher would be disappointed if any of her students came to believe that

---

[3] Our formulation thus far leaves it open whether the apparent reflexivity is eliminable. Grice took it to be, whereas I will argue below that it is not.

'Moon Over Miami'

Performing an action with (a) the intention to produce a belief in an audience, (b) the intention that this effect be achieved at least in part by the audience's recognition of the speaker's intention, and (c) the the intention that the audience be aware of her intentions (a) and (b)

The River Rat: S takes into the house that A is thinking of buying and lets loose a big fat river rat. He knows that A is watching him and knows that A believes that S is unaware that he, A, is watching him.

Performing an action with (a) the intention to produce an effect on an audience, (b) the intention that this very intention be recognized by that audience, and also (c) the intention that this effect on the audience be produced at least in part by their recognition of the speaker's intention.

Herod shows Salome St. John's head on a charger with the overt intention of making Salome believe that St. John is dead; he intends that she discern that St. John is dead and also that this very intention of his be recognized.

Performing an action with the intention of influencing someone's beliefs, while intending that this very intention be recognized

I leave Smith's handkerchief at the scene of the crime with the intention of making the police believe Smith is the culprit. Still, it's not plausible that I mean that Smith is the culprit.

Performing an action with an intention, successful or not, of influencing someone's beliefs

In putting on a coat I might lead an observer to conclude that I am going for a walk; yet this is not a case of speaker meaning.

Performing an action that influences the beliefs of an observer

Figure 3.2: Grice's Ladder

theorem in part because she intended them to. Rather, she wants them to believe the theorem on the strength of the proof that has been given. Reflexive communicative intentions are out of place here, yet speaker meaning remains.

Reflexive communicative intentions are not necessary for speaker meaning. In fact, speaker meaning can occur without a speaker intending to produce *any* beliefs in an audience. A framed suspect might mean that she is innocent in saying, "I am innocent!", yet be fully aware that no one will believe her and perhaps, being realistic, not intending to convince anyone. She might not even intend her interrogators to believe that *she* believes she is innocent, since she might know that they are certain she is lying. Or, gazing into my newborn daughter's eyes I might say, "All things valuable

are difficult as they are rare,'' meaning what I say, without having the slightest intention to produce beliefs or other attitudes in her or in anyone else. Again, in the film *Sleeper*, Woody Allen's character Miles Monroe comes across, while exploring alone, a genetically modified chicken the size of a small house. Miles remarks, "That's a big chicken." In saying this he does not seem to be intending to produce an effect on anyone, himself included.[4]

Following a suggestion of Schiffer (1972, p. 15), Strawson (1970, p. 7), and Bennett (1976, p. 271), Avramides proposes in response to these kinds of case that the speaker is addressing himself, intending in particular to produce a certain cognitive effect in himself.[5] While it may be that in the newborn daughter and *Sleeper* cases the speaker is addressing himself, it neither follows, nor does it seem true, that in those cases the speaker is intending to produce any cognitive effect in himself. Certainly we sometimes address ourselves in order to produce a cognitive effect: 'I can do it!' as I sprint up the steep road, or '945–6743, 945–6743' as I try to internalize a phone number I just got out of the phone book. However, I already believe that all things valuable are difficult as they are rare. In fact it is a belief I have held firmly since encountering it in Spinoza two decades ago, and I actively believe it as I reflect upon the number of diapers I will have changed by the time I am forty. As a result it is quite unclear what cognitive effect I might be trying to produce in myself in saying what I do. Again, Miles Monroe does not need to produce in himself, or strengthen

---

[4] This discussion is indebted to Davis 1992a. We observe here also that Armstrong 1971 quite reasonably offers an account of speaker meaning in terms of objectives rather than intentions, his reason being that the latter notion is narrower than the former. One who intends a certain result must believe that the thing aimed at is within her power, while one who has that result as an objective need not do so. Presently we shall show an affinity between Armstrong's position and that offered below. However, just replacing 'intention' with 'objective' in Grice's account will not deal with the cases we have considered. It is not part of my objective to produce an effect in my newborn daughter in uttering the last line of Spinoza's *Ethics*. Similarly, it need be no part of the objective of the framed suspect in maintaining her innocence to produce effects on her interrogators. Instead she may say what she does in order to make public, for anyone who may be concerned with the matter, her avowal of innocence. Her objective is simply to establish a pattern of consistently maintained innocence.

[5] Avramides' discussion here is confusing because she first responds to a case, due to Harman, of a person maintaining a proposition in full knowledge that no one will believe him, with the words, 'I think that in Harman's case the speaker is not really speaking to an audience at all' (p. 64). But then two pages later Avramides writes, 'The misleading thing about Harman's case is that there appears to be an audience present. The speaker, however, does not really address his utterance to those present.… If this is true, why not say that in such cases the speaker intends his audience to be himself…' (p. 66). I shall take Avramides to hold the view that in these cases the speaker *does* have an audience, namely himself.

or activate, the belief that the chicken before him is big. His eyes have already done that for him. Likewise it is far from clear what belief the framed suspect might be trying to produce or strengthen or activate in herself as she maintains her innocence. The suspect knows perfectly well that she has never set foot in the part of town in which the crime was committed, and that she has no idea how to use the garotte with which the victim was killed.

It might, alternatively, be suggested that in the above cases the speaker addresses a virtual audience. This idea can take one of two forms. First of all, it might be suggested that the speaker is imagining addressing someone, not unlike the way in which a child might address an imaginary friend. Yet I certainly don't feel as if I am addressing an imaginary person when I utter the Spinozistic dictum, and I doubt that the framed suspect feels this way either. We can likewise easily construe Miles Monroe as not imagining that he is addressing anyone. On the other hand, the "virtual audience" proposal might take the form of a suggestion that in these cases, it is the speaker's intention that were someone in a position to understand and appreciate his utterance, the remark would produce an effect on that audience, for instance belief (Hyslop 1977). This proposal may well be correct, and we need not take issue with it. The reason is that one can have this counterfactual intention without intending to produce any effects on an audience. It is consistent to think both (a) were someone in a position to understand and appreciate an utterance, it would have an effect on them, and (b) no one is in a position to understand and appreciate the utterance. Likewise, one could (a) refrain from intending one's utterance to have an effect on anyone, and (b) intend that were someone in a position to understand and appreciate one's utterance, the remark would produce an effect on them. A counterfactual recasting of the notion of audience-directed intentions does not salvage the idea that speaker meaning requires audience-directed intentions.

We have shown difficulties in each of two ways of defending the idea that speaker meaning requires intentions to produce effects on an audience. Our responses to these defenses might raise the question why in the cases we have considered the speakers said anything at all rather than keeping their mouths shut. In the interrogation example the answer, as we saw above, is simply that the suspect wants it on record that she has maintained

her innocence. In the *Sleeper* example the answer is also clear: sometimes we are struck with a thought that it is natural to express. This is not different in principle from the way in which we are possessed by an emotion that it is natural to express. Indeed the two phenomena often overlap. In the *Sleeper* example, the speaker is not only expressing his belief that the chicken is large, but seems also to be expressing a sense of awe or wonder at the enormity of the fowl. My utterance of the last line of Spinoza's *Ethics* expresses my sense of both the burdens and pleasures of parenting, and I can find satisfaction in thus expressing myself. One need look for no deeper explanation than this.[6]

Speaker meaning does not require intentions to produce beliefs in an audience, even oneself, to say nothing of intentions to produce beliefs by means of recognition of one's intention. What then, *is* speaker meaning, and how could Grice and others have gone so wrong in defining it in terms of audience-directed intentions?

3.4. Intention: nests and hierarchies

After Grice formulated his reasons for imputing reflexive communicative intentions to producers of speaker meaning, among philosophers a cottage industry sprang up concerned with whether even these intentions were sufficient to produce the phenomenon in question. Those convinced that reflexive communicative intentions are necessary for speaker meaning soon agreed that they are not sufficient in light of cases such as one devised by Strawson 1964:

The River Rat: Homebuyer is inspecting a house for possible purchase, and his friend—call him Friend—is concerned to convince him that the home is rat infested. Friend arrives at the house at a time when he knows that Homebuyer is inside, and although he knows he is being watched, skulks around to make Homebuyer think that he, Friend, believes he is acting unobserved. Friend has a river rat that he places in a salient position for Homebuyer to see. He intends for Homebuyer to see the rat and reason as follows: "Although the rat display was rigged, Friend would not have put it there unless he believed that the house really

---

[6] The satisfaction found in these cases of self-expression is over and above the well-documented way in which expression of emotions can have a therapeutic effect, often with positive repercussions for bodily health. For further discussion see Pennebaker 1990.

is rat infested; hence Friend, who is reliable and honest, must intend me to believe that the house is rat infested.''

Friend thus has the following complex of intentions. In so placing the rat, he intends Homebuyer to believe that the house is rat infested; further, he intends to produce this effect in Homebuyer by means of his recognition of Friend's intention. Friend thus meets Grice's conditions for reflexive communicative intentions. It's also colloquial enough to describe Friend as trying to convey to Homebuyer that the house is rat-infested. However, it does not seem that in placing the rat under the conditions we have described, Friend means that the house is rat infested.

Some proponents of Grice's program of construing speaker meaning in terms of communicative intentions responded to this case by adding another clause to the definition of speaker meaning.[7] In addition to (a) the intention to produce a belief in an audience, and (b) the intention that this effect be achieved at least in part by the audience's recognition of the speaker's intention, we require further that the speaker intend (c) that her audience be aware of her intentions (a) and (b). Unfortunately, examples that satisfy these three clauses, without, however, seeming to be cases of speaker meaning, have been found.[8] Trying to rule out these further cases with a more complex analysis threatens to make the Gricean approach to speaker meaning psychologically unrealistic.

Other authors have suggested that what unites the various cases we have observed to make trouble for the Gricean approach is that they involve a kind of covertness, or at least a lack of overtness. What is missing in the handkerchief case is the intention, on the part of the man planting the handkerchief, that his intentions be publicly available. Similarly, part of the intent of Friend in the River Rat case is to be covert (he wants Homebuyer to be unaware that he knows he is watching him), and this is what seems to prevent that case from being one of speaker meaning in spite of the complex communicative intentions involved.

However, the foregoing subtleties might seem moot in light of our realization that speaker meaning does not require intentions to produce beliefs or other attitudes. What at first seemed an ascent of Grice's ladder may now appear a fool's errand. Did we get off on the wrong foot in judging

---

[7] For a thorough review of the literature in the wake of Strawson, see Vlach 1981.
[8] One such case is the ''Moon Over Miami'' example of Schiffer 1972.

that, since the overcoat and handkerchief examples do not exemplify speaker meaning, some more complex congeries of intentions must do the job? I suggest that the search for ever more complex intentions such as would be sufficient for speaker meaning conflated two things: (1) intentions to produce cognitive effects in some audience, and (2) intentions that one's intentional state be *manifest*, that is, publicly accessible, but not necessarily in fact discerned by anyone. One can have the latter intention while harboring no intention whatever that anyone come to be aware of one's own state of mind. Instead one need only intend that the intentional state be "out there": there to be discerned by anyone concerned to look whether or not anyone ever does.[9]

Intentions to produce beliefs in an audience are of course commonly found in the company of speaker meaning. Our discussion thus far shows that nevertheless it can manage without them. What it cannot eschew is overtness. Recall the above case in which a person dons an overcoat, leading others to conclude that she plans to go for a walk. Here there is no speaker meaning. However, imagine a revision of the case in which Hermione is arguing with a group of friends about the weather. Everyone else has been warning her of the dangers of going outdoors in the storm. She now dramatically swings open the closet door, dons an overcoat with a flourish, and grabs an umbrella while beadily meeting the gaze of her friends. Here it seems clear that she means that she is intending to go for a walk, and I suggest that what differentiates this from the earlier case in which someone's intention to go for a walk is discernible is that Hermione is making her peripatetic intention overt.[10]

---

[9] There are two notions of reason, one requiring, the other not requiring, a subject's awareness of the reason in question. In accord with the latter notion, I might, unbeknownst to me, have a reason to believe that one of my employees is embezzling funds even if I am not aware of that reason because I have not noticed the Byzantine character of the balance sheets he has submitted. Reading a proposal of Armstrong's according to this "objective" notion of a reason allows us to see an affinity between it and the approach taken here. Armstrong suggests (1971, p. 435) that a speaker S who means that P has the objective that the audience should have reason to believe that S believes that P. On the objective notion of a reason, S's objective need not include producing an effect on any audience. It need only include producing evidence, there for anyone to see, that S has a certain belief.

[10] Grice 1957 considers a case that is superficially similar to this one, but which is less clearly an instance of speaker meaning. In that example a father leaves lying around on the floor for the mother to see the bits of the china that their child has broken. Here it is doubtful whether the father means anything in leaving the china on the floor. As Grice describes the case, the mother might be open to any of a variety of reasons why the china is on the floor: perhaps the father didn't notice it, or had other more pressing issues to attend to, or was too furious to clean it up. This is why the example

What, then, is overtness? To be overt an action of mine need not actually be discerned by anyone; I could flagrantly show contempt for someone, and if they are too distracted or obtuse to recognize what I am doing this will make the action no less overt. Intending that an action of mine be overt should likewise not require intentions to produce effects on others. Instead, it must, very roughly, be an intention that everything of relevance be "out in the open". One way of articulating this thought is with a hierarchy of conditions having the following structure:

(a) I intend to show my contempt,
(b) I intend that my intention (a) be manifest,
(c) I intend that my intention (b) be manifest,
and so on.

It is, however, difficult to accept this as a description of what is going on when I, for instance, flagrantly show my contempt for someone. The reason is that it is unlikely that I would harbor such an unending hierarchy of intentions. How then may we articulate the wanted notion of overtness? Another approach capturing the wanted notion of overtness found in an important class of cases of speaker meaning is in terms of an action done intending that (a) something be publicly discernible, and (b) this intention itself be publicly discernible as well. This is an intention part of whose content, the part expressed in clause (b), refers, inter alia, to the very intention whose content it is, with the result that there is no finite way of writing down the content of this intention from which self-reference has been elided. That does not, however, imply that to grasp this intention a thinker must grasp a thought that is in any sense infinitely "long". Instead the thought can be expressed, and thus grasped, quite succinctly.[11] Consider an analogy. I might be regaining consciousness from an accident that, as it was about to occur, seemed certain to kill me. As I contemplate my good fortune before opening my eyes to assess the damage I might think, 'This

is only superficially similar to the case of Hermione in the text. Hermione leaves little doubt about the intention with which she is acting. This is to say that Hermione overtly displays her intention, whereas the father does not make overt the intentions with which the china has been left lying around. Accordingly, while both the china and Hermione examples involve showing rather than telling, the awkwardness of describing the former example as a case of speaker meaning does not undermine the contention that Hermione means that she intends to go for a walk.

[11] Here I am indebted to Harman 1974, 1977, and to Recanati 1986.

thinking is miraculous.'[12] The content of this thought refers to the thought token (a particular thinking with a spatiotemporal location, or at least a temporal location, and a content), and says of it that it is miraculous. It will, then, be true just in case that very thought token is miraculous. The fact that the content of this thought token cannot be finitely rewritten in a way that eschews all reference to the thought token itself, does not upset the fact that this content has determinate truth conditions, and a determinate meaning.

"Self-referential" thoughts do not, as such, appear incoherent.[13] If their content can be anchored by a particular thinking, perceiving, intending, and so on, there need be no difficulty in principle in grasping that content. This point is applicable to our elucidation of speaker meaning, for as we have seen, one way of meaning that P, namely that associated with (but not exhaustive of) propositional speaker meaning, is overtly to make the fact that P manifest. That is, in such a case one intentionally makes P manifest, intending as well that this very intention to make P manifest, itself be manifest. That suggests an elucidation of one form of speaker meaning along the following lines:

> *Factual Speaker Meaning*: Where P is an actual state of affairs, S factually speaker-means that P iff
> 1. S performs an action A intending that
> 2. in performing A, it be manifest that P, and that it be manifest that S intends that (2).

P might be the fact that it is windy outside, or that S believes that it will rain, with A being the uttering of certain words or a non-conventional action such as the throwing open of curtains to reveal a looming storm. In the Herod case the P in question is the fact that St John is dead, and the action is Herod's presenting of St John's head on the charger. In the case of an irritated person scowling intentionally, the P in question is that he is irritated, and the A is the irritated person's production of the angry scowl. As with the case of the thought, 'This thinking is miraculous', Factual Speaker Meaning requires, in condition (1), that an intention (as opposed

---

[12] The example is due to Peacocke 2005.

[13] Some authors, such as Harman, hold that all intentions are self-referential. We need not commit ourselves to this strong claim in order to hold that intentions germane to speaker meaning are self-referential.

to a thought) be tokened. The content of that intention is then articulated in condition (2), which refers, inter alia, to the intention-token whose content it is. That intention-token will, in turn, be satisfied (as opposed to being true) just in case condition (2) holds. Once again, P might be manifest without anyone being aware of this fact. As a result one can intend that P be manifest without intending to produce effects on others.[14] In light of our account of Factual Speaker Meaning, then, one can mean something without intending to produce effects on any audience, and one can mean something in the course of making some state of affairs manifest.

Speaker meaning, as we have seen, also comes in an objectual variety, in which an agent means an object rather than that something is so. This has been exemplified in such cases as a person referring to someone with words, as in, 'Look over there!', or without, as in their simply pointing to or gazing pointedly at the thing in question. After observing such an act we may ask, "Which one do you mean?" if we are still unsure. Such objectual meanings must, like their factual analogues, be overt, and we now know how to elucidate this requirement. Where α is a perceptible object,

Objectual Speaker Meaning: S objectually speaker-means α iff
1. S performs an action A intending
2. α to be manifest, and for it to be manifest that s/he intends (2).

Instead of the factual construal offered above, Herod's action may be construed as done with the intention of making St John's severed head manifest, while intending that this very intention be manifest as well. In the case of the overt scowl, the α in question might be the agent's anger,

---

[14] Some authors have been skeptical of the very possibility of self-referential intentions. (Such skepticism is not to be confused with skepticism of the doctrine that all intentions are self-referential. One may accept that self-referential intentions are possible while remaining neutral on the question whether all intentions are self-referential.) Thus for instance Seibel 2003 writes, "the content of [the self-referential intention] contains an element which refers to the intention itself. But what does that element look like? … How does it single out the intention and nothing but it? By identifying features, i.e., properties which are exclusively possessed by the intention? But what could be these features?" Intentions admit of the same act/object dichotomy as do many other intentional states, and like other mental events, are spatiotemporally located. On the modest assumption that no two intentions have identical spatiotemporal coordinates, we may then use such coordinates to individuate intentions. One might still wonder what the *content* is of an intention whose content refers to that very intention, which itself comprises both an intending (a state or act) and a content. One answer may be given in terms of an analogue of truth conditions applicable to intentions, namely satisfaction conditions. Just as the thought, 'This thinking is miraculous' will be true just in case that thinking is, indeed, miraculous, so too, the intention, 'This intention shall be manifest' will be satisfied just in case that intention is, indeed, manifest.

the agent intending that his intention to make his anger manifest be itself manifest. In fact, particularly when no conventional language is used, it may be difficult to know whether the case is objectual or factual speaker meaning. When an indicative sentence is uttered, that is strong evidence that the speaker's meaning is not objectual. Failing that evidence, however, not only may it be difficult to tell, we need not assume that there must be a fact of the matter whether the speaker's meaning is factual or objectual. It may well, that is, be indeterminate whether the speaker's meaning is one or the other. If that is so, then in such cases, a choice between an objectual and a factual way of representing what the speaker means, if such a choice must be made, is presumably to be made on pragmatic grounds.[15]

Both Factual Speaker Meaning and Objectual Speaker Meaning pertain to cases in which the speaker takes what she means to obtain (for the factual case) or exist (for the objectual case). Otherwise it is difficult to see how she could intend to make what she means to be manifest. Clearly we can also mean propositions that we do not think are true, as when we lie, and evidently we can also mean objects that we do not believe to exist, as in make-believe. We can also mean propositions whose truth we are not in a position to demonstrate. These are reasons why the above conditions are not formulated as necessary for speaker meaning generally. If our exploration of speaker meaning in this chapter has been accurate, however, it would be a mistake to infer, from the premise that *some* forms of propositional speaker meaning do not involve intentions to manifest facts, to the conclusion that *none* ever do. Likewise it would be a mistake to infer, from the premise that some forms of objectual speaker meaning do not involve intentions to display or manifest objects, to the conclusion that none ever do. We are now, however, in a position to build on what we have learned to develop a general account of speaker meaning.

3.5. Speech acts and handicaps

Factual speaker meaning usually occurs in the context of a speech act. Let me make clear that speech acts are to be distinguished from acts of

---

[15] In Chapter 6 we will return to issues of indeterminacy in ascribing attitudes to agents.

speech. When I test a microphone, utter lines on stage, or practice a speech in the shower, I am performing acts of speech but no speech acts. By contrast, a speech act is any act that can be performed by, under the right circumstances, speaker-meaning that one is doing so. I can raise a question ('I ask you what time it is'), make a statement ('I state that it is 5 p.m.'), issue a command ('I command you to make the appointment on time') by saying that I am doing so in such a way as to speaker-mean it. This is why questioning, stating, and commanding are speech acts. I can also perform one of these acts by speaker-meaning a content in a certain way but without saying that I am doing so. I can assert that snow is white without saying that I am asserting that snow is white. I can do that simply by uttering the words, 'Snow is white', meaning this as an assertion.

Since I cannot scare or persuade you by saying that I am doing so, scaring and persuading are not speech acts. Although speech acts are only a small portion of the acts that we are able to perform, some such acts are nevertheless quite momentous: so long as I have the authority and other conditions are in place, I can excommunicate, appoint, bequeath, or even declare war by speaker-meaning that I am doing so. In this section I will argue that central features of certain speech acts can be explained in terms of their being handicaps in the sense of that term introduced in Chapter 1.

How can the biologist's notion of a handicap be so much as relevant to the elucidation of speech acts? To see why it is, consider that it's a bit of common sense that speech acts often, though not always, have expressive dimensions. It is essential to a wide variety of speech acts that when performed in all propriety they express a state of thought or feeling.[16] A sincere assertion expresses one's belief, a sincere promise one's intention, a sincere apology one's regret. Further, one of the dicta we formulated in Chapter 2 has it that expression is a species of showing; another has it that it is a species of showing *one's* state of mind, state of heart, or state of experience. It follows that one who sincerely asserts P shows her belief that P; one who sincerely promises to do something shows her intention of doing so, and so forth. We have also suggested, in Section 3.2, that this notion of showing comes in at least three forms: showing that something is so, showing α (where α is an object of

---

[16] Green forthcoming a develops the point in further detail.

perception), and showing how something feels, be it an experience, emotion, or mood. Beliefs, intentions, and the like are not, apparently, possible objects of perception, nor is it plausible that there is a way that a belief or intention feels. If we are to show our belief by means of a sincere speech act, it will evidently have to be by showing that we harbor the belief that we have expressed. Yet how can a mere speech act provide enough evidence for the presence of a belief in order to show it?

Consider a well known conception of assertion. In *Speech Acts*, John Searle characterizes assertion of P as, "an undertaking to the effect that P represents an actual state of affairs" (1969, p. 66). Because undertakings are more naturally thought of as undertakings to action than as having propositional objects, it is not entirely clear what such an undertaking consists in. Minimally, we may construe an undertaking to the effect that P represents an actual state of affairs as putting the assertor in a position such that she is correct on the issue of P if P obtains, and incorrect on the issue of P otherwise. Assertion, however, does more than put the assertor at risk of being either correct or incorrect on the issue of the asserted proposition. After all, one who guesses that P is also correct on the issue of P if P obtains, and incorrect on that issue otherwise. Instead, when a speaker asserts that P, it is within the rights of an addressee to reply with the challenge, "How do you know?" In response to such a challenge the assertor should either offer reasons of her own, or defer to another ostensibly reliable authority ("I read it in the *Times*," "Susan told me."). If she cannot do either of these things she should retract the challenged assertion.[17]

Compare assertion with the case of conjecture: If I offer P as a conjecture, then it is inappropriate to reply with the challenge, "How do you know that?" For instance, I might conjecture that a black hole inhabits the center of the Milky Way. As with the case of assertion, what I say is right or wrong on the issue of P depending on whether P is true. However, it is inappropriate to reply to my black hole conjecture with a challenge to show that I know what I say to be true. A legitimate challenge to a conjecture would instead come in the form of showing that what I say

---

[17] We need not assume that the normative notions invoked here are reducible to moral norms or to norms of theoretical or practical rationality. They may be, but we need not settle that issue here.

is demonstrably wrong, or at least very unlikely. In that case I would be obliged to retract that conjecture. Likewise, a conjecture should be backed with some justification or other; failing that it should be put forth as a guess rather than a conjecture. However, if you do put it forth as a conjecture, the required justification need not be as compelling as that for assertion.

Assertions, conjectures, suggestions, guesses, presumptions, and the like are cousins sharing the property of commitment to a propositional content. They differ from one another in the norms by which they are governed. In addition, in assertions, conjectures, suggestions, and presumptions, the speaker incurs a vulnerability—not just a liability to being in error, but also a mandate to defend what she has said if appropriately challenged. These liabilities to error and injunctions to defend what has been said put the speaker at risk of losing credibility in the community in which she has a reputation.

Assertions, conjectures, suggestions, and presumptions, and the like thus carry a cost. Why should speakers bother to pay it? Presumably speakers are willing to incur such costs because the information they contribute to the group can help achieve the group's aims—finding food, prevailing over rivals, securing shelter, and so forth. In addition, an individual might want to be the first one to provide such information in order to enhance her or his credibility within the group: saying something unobvious and useful to the community can enhance one's prestige. Yet in the absence of some mechanism for vouchsafing honesty, these speech acts will be prone to abuse by those who take assertion lightly in the hope of scoring some epistemic points. After all, if I say enough, I'll eventually get something right. This temptation threatens to make assertors, conjecturers, and the like on the whole less worthy of our belief: liars, and those who say things on insufficient evidence, threaten to undermine the efforts of sincere speakers in the way that brightly and dishonestly colored frogs threaten the credibility of signals sent by those frogs who are both brightly colored and noxious. In chapter 2 of *Utilitarianism*, Mill emphasizes the value of making statements when he refers to, "the trustworthiness of human assertion, which is not only the principal support of all present social well-being, but the insufficiency of which does more than any one thing that can be named to keep back civilization, virtue, everything on which human happiness on the largest scale depends" (Mill 2002, p. 23). That's pretty high praise.

Mill may be overstating the point, but he is responding to a legitimate anxiety.

I propose that many linguistic communities have found ways of circumventing this danger of the erosion of speakers' credibility. Recall that the long eye stalks of the male stalk-eyed fly show its viability. The magnitude of the spider's vibration on its web show its size. So too, the liabilities that an assertor undertakes by incurring the commitments characteristic of that speech act give strong evidence that the assertion is both sincere and justified. Certainly not conclusive evidence: many lie, and many believe things on insufficient evidence. Nevertheless, by asserting something sincerely you perform an act that, because of the handicap it involves, *shows* your belief. So, however, does a sincere conjecture. In addition, the assertion, precisely because of the strict standards to which it is held, shows your belief *as justified at a level appropriate for knowledge*, while the sincere conjecture does not. (The sincere conjecture shows your belief as having some justification, but not sufficient justification for knowledge.)

We have discussed speaker meaning in objectual and factual forms. However, when I mean that P as an assertion (rather than as a conjecture, etc.) I do not therefore mean that I am asserting that P. I might just mean P. Yet, the fact that I am putting forth P as an assertion rather than a conjecture is part of what I mean. This is why it would be appropriate for someone who doesn't know the force of my utterance, to ask, "Do you mean that as an assertion, a conjecture, a guess, or what?" How, then, may we understand the contribution that illocutionary force makes to speaker meaning? We may do so as follows. I can mean P *as an assertion*, rather than in some other way. What makes it the case that I mean P as an assertion is that in so doing I undertake a certain set of responsibilities and liabilities, namely those that we have outlined above as distinctive of assertion as against other speech acts. What makes it the case that I mean P as a conjecture is that in so doing I undertake a distinct set of responsibilities and liabilities. And so on. My meaning P with one force rather than another is not a matter of *what* I mean but rather *how* I mean it. We are now in a position to offer a characterization of speaker meaning distinctive of speech acts with propositional contents:

Illocutionary Speaker Meaning: S illocutionarily speaker-means that P φ'ly, where φ is an illocutionary force, iff

1. S performs an action A intending that
2. in performing A, it be manifest that S is committed to P under force φ, and that it be manifest that S intends that (2).

One illocutionarily speaker-means that P assertorically if one performs an action intending that in so doing it be manifest that one is committed to P assertorically, and that this itself be manifest. One illocutionarily speaker-means that P as a conjecture if one performs an action intending that in so doing it be manifest that one is committed to P as a conjecture, and that this itself be manifest. These characterizations are not circular because 'illocutionarily speaker-meaning P as an assertion', 'illocutionarily speaker-meaning P as a conjecture', and so on may be fleshed out in terms of conversational norms. One way to make manifest that you are committed to P in a certain way is to make explicit that you are: use such words as ''I assert that P.'' Parenthetical expressions work in this way as well, as in ''P, I claim.'' Or you could let context make clear the force of your utterance, so that by uttering P free-standing, rather than embedding it in a larger sentence, you rely on context and the reasoning capacities of your addressees to discern that P is meant as an assertion rather than in some other way.[18]

Speaker meaning is a matter of overtly showing an object, overtly showing a state of affairs, or overtly showing one's commitment—both the modality of that commitment and its content. Each of the three conceptions of speaker meaning we have offered—Objectual Speaker Meaning, Factual Speaker Meaning, and Illocutionary Speaker Meaning—corresponds to one of these three forms of speaker meaning. For this reason we are now in a position to give a general account of speaker meaning, including the propositional, objectual and illocutionary cases:

---

[18] This definition of illocutionary speaker meaning only applies to speech acts with propositional contents. It does not apply to speech acts with interrogative contents (*where is John?, how many apples are in the bowl?*), or what I would call imperatival contents (*to shut the door, to eat the apples*). I believe that the account can be generalized to include such cases, but such a generalization would over-burden this chapter.

Also, just to forestall any worries about an impending regress: One who asserts P by saying that she is doing so most likely will utter some such words as, ''I assert that P.'' What, then, is the status of the embedding sentence? It too will be being put forth with assertoric force. You could of course indicate that by saying that you are doing so, but this approaches the margins of preciosity. Most likely you will just let context make clear that you are putting forth the sentence, 'I assert that P' as an assertion rather than in some other way.

Speaker Meaning: S speaker-means something just in case S either objectually speaker-means something, factually speaker-means something, or illocutionarily speaker-means something.

If the foregoing is correct, then we may understand speaker meaning as a species of signaling, whether or not that speaker meaning is sincere or justified. This framework will also enable us to see culture, human or otherwise, as standing in where biological fixity does not to create stable signaling systems. This is not to say that all cultures are sensitive to distinctions among asserting, conjecturing, guessing, and the like; some might well not bother with these subtleties. I do, however, suggest as an empirical hypothesis that in those cultures in which one can handicap oneself by means of a speech act, communicative stability is served.

3.6. Alternative accounts of speaker meaning

In a survey of literature up to 1980, Vlach 1981 shows that the theories propounded in Grice 1957, Grice 1969, Armstrong 1971, and Bennett 1976 founder on one or another version of the "proof" example, in which the geometry teacher intends her students to believe what she says, but not because of their recognition of her intention that they do so.[19] Vlach also shows that many more recent analyses of speaker meaning are objectionable for various reasons. Following a suggestion of Searle's, Vlach propounds an account of speaker meaning as follows: U means P by x iff U does x in the belief that he is thereby committing himself to P (1981, p. 382). This suggestion has the virtue of allowing for the possibility of speaker meaning in the absence of intentions to produce effects on an audience. It also captures a crucial feature of a significant class of speech acts, namely that they involve a form of commitment. However, the undertaking of commitment is not a necessary condition of every speech act. Speech acts are typically instances of speaker meaning,[20] but not all

---

[19] See, however, note 3 above for a qualification of this remark about Armstrong.

[20] I say typically here because there could be speech acts governed by a "strict liability" constraint. Thus for instance grasping a certain conventionally defined object could constitute agreement to enter into military service for a certain country. Arguably, this was precisely the case in England in the

involve the undertaking of commitment. For instance, it is hard to see how imperatives such as commands involve the undertaking of commitment. Likewise for greetings such as ''Hello'' or other of Austin's ''behabitives'' such as apologies and congratulations. Furthermore, while it seems clear that conversational implicature is a form of speaker meaning, one who conversationally implicates something is not thereby committed to the content thus implicated. If in answer to the question, 'Where is Mary?' asked by someone hoping to visit her, I answer, 'Somewhere in France', I may conversationally implicate that I am unable to be any more specific about her whereabouts. However, in so answering I have not committed myself to the proposition that I cannot be any more informative than I have been. This is in contrast to conventional implicature, in which a speaker's use of an expression commits her to a content by virtue of the conventional meaning of that expression even though that content is no part, or at least not an explicit part, of what she says.

Neale 1992 also denies that reflexive intentions to produce beliefs in an audience are necessary for speaker meaning, while adhering to the view that speaker meaning admits of a uniform characterization. He suggests the following account of speaker meaning: by uttering x, U meant that P iff for some audience A, (1) U uttered x intending A actively to entertain the thought that P (or the thought that U believes P), (2) U uttered x intending A to recognize that U intends A actively to entertain the thought that P, (3) U does not intend A to be deceived about U's intentions (1) and (2). Observe first that this is a considerable weakening of the analysis, since it will follow from this account that under normal conditions, one who asserts ''If P, then Q'' will mean P and Q, and not just the conditional. Similarly, normally a person who asserts ''It is not the case that P'' will mean both P and its negation on Neale's account of speaker meaning, and one who recites P on stage will mean P. These consequences do not as such undermine Neale's proposal. However, that proposal runs afoul of the fact that we sometimes direct our remarks to those whom we know are not even comprehending what we say. We have already seen this in the

eighteenth century: grasping the ''King's shilling'', whether one intended to or not, and whether or not one was aware of the consequences, constituted an agreement to enter into England's navy. That agreement might have been a speech act in spite of the fact that in so grasping the Queen's shilling one does not mean that one is willing to join her navy. See Hare 1989 and Green 1997 for further discussion, and I recur to the point in Section 4.3.

examples of my speaking to my newborn daughter, and of Miles Monroe remarking on the size of the chicken. Or I might, out of determination to fulfill the duties of my contract, carry through a complex lecture on a topic that I know is well over the heads of my students. Here when I state that the class S of all Borel sets coincides with the ó-ring generated by the class U of all open sets, I exemplify speaker meaning. However, I do not meet Neale's condition (2), for I do not intend my audience to entertain the thought that the class S of all Borel sets coincides with the ó-ring generated by the class U of all open sets. Indeed, because I am sure that none of my students is listening, I do not even intend that they entertain the thought that I believe this. As a result Neale, who waters down the Gricean requirement of reflexive communicative intentions about as much as possible while retaining its spirit, does not capture what is central to speaker meaning.

Davis (1992a, 1992b, 2003), as we have mentioned, is one source of our earlier arguments that speaker meaning does not require audience-directed intentions. His arguments apply to well known theories offered by Grice 1957, 1969, 1982, Schiffer 1972, and Avramides 1989, among others, and I shall not rehearse them here. His own position, developed most fully in Davis 2003, naturally provides an account of speaker meaning that eschews reflexive communicative intentions: S meant that P by producing e iff S performed an observable action as a direct and undisguised indication that he occurrently believes P (p. 57). 'Observable', because Davis holds that all speaker meaning requires making one's thought, or ostensible thought, public. 'Direct' because Davis wishes to focus his discussion on what he calls the 'exclusive' sense of 'meaning', as opposed to an 'inclusive' sense which includes all that one implies, conversationally, conventionally, or otherwise. Undisguised because of the following example: John and Mary are trying to fool George, and so Mary pretends that she is in great pain. John rushes to her aid, performing an observable action as a direct indication that he believes she is in pain. However, Davis observes, John does not here mean that Mary is in pain; instead he is merely pretending unintentionally to manifest that belief. Davis refers to this as a covert simulation of an unintentional manifestation of belief, and for brevity rules out cases of this sort by requiring that speaker meaning must be undisguised (ibid, pp. 54–6). Finally, 'occurrent' because Davis holds that one can only mean P if P expresses an occurrent thought.

One who asserts P expresses the belief that P, indeed expresses *her* belief that P if she is sincere, and so Davis gives us a necessary condition for assertion. Yet even leaving aside the non-indicative case, we can see that the condition is not necessary even for other indicative speech acts such as supposition or conjecture. Given his remarks about guessing (1992b, p. 226) Davis would likely treat one who puts forth P as a conjecture as expressing the belief that P is probably true. But that is not what I mean when I put forth P as a conjecture. Granted, I would not be inclined to conjecture that P if I did not think that P was at least reasonably likely, but this does not imply that this is what I mean when I utter, 'The cause of the perturbations in the planet's orbit is a distant white dwarf,' with the force of a conjecture. The difference comes out dramatically when we understand the kind of commitment undertaken in the forwarding of a conjecture. One who conjectures that P is apt to be right about P if indeed P is true; apt to be wrong about P if P is not true. This much cannot be said for one who means or believes that P is probably true. I can mean or believe that P is probably true without being shown wrong if indeed P turns out to be false. Probabilistic assessments do involve the undertaking of commitments, but the sorts of commitments they involve are quite different from those undertaken by such "qualitative" judgments as conjectures.[21] Davis's treatment also founders on examples of putting forth P with the force of a supposition for the sake of argument; in such a case the speaker means P, but she expresses neither belief nor any particular degree of belief that P.

Davis's account of speaker meaning requires that the thought in question be occurrent, but this is also too restrictive. Many things that we say, and mean, are said out of habit; in other cases things we mean are conceptually quite complex, and when we say them it is not clear that we need to entertain the thought expressed by the words that we use in order for our utterance to be a case of speaker meaning. I recently reproduced the proof that the class S of all Borel sets coincides with the ό-ring generated by the class U of all open sets, and just last Tuesday I was rehearsing both the proof and its conclusion. I still believe that conclusion, and I could formulate it verbally as well. Suppose I do formulate it verbally without consciously contemplating the theorem. This seems a clear case of speaker meaning in

---

[21] Green 1999c defends this point in further detail.

spite of the fact that right now I lack the energy or concentration needed to entertain the thought that those words express.[22]

We may also note that not all speech acts characteristically express intentional states. In discussing Vlach's proposal we have already mentioned greetings. Austin's ''exercitives'' are another case: it is difficult to see that excommunicating, for instance, expresses any intentional state. Davis's proposal is thus at best a sufficient condition for speaker meaning, and is therefore in no conflict with the sufficient condition that we have offered above in the form of our Sufficient Condition for Propositional Speaker Meaning.

Sperber and Wilson (1995) offer an account of communication with affinities to the view of speaker meaning propounded here. In particular, while Sperber and Wilson do not attempt to characterize speaker meaning, they do offer a view of communicative intentions from which a view of speaker meaning may be developed. On their official account, a fact is *manifest* to an individual just in case that individual is capable of either perceiving it or inferring it from other things she knows or thinks probable (p. 39).[23] A *cognitive environment* of an individual is the set of all facts that are manifest to him (p. 39). A cognitive environment is *shared* just in case it is manifest to more than one individual (p. 41). A *mutual cognitive environment* is a shared cognitive environment in which it is manifest which people share it (p. 41). Every fact or assumption in a mutual cognitive environment is what Sperber and Wilson call *mutually manifest* (p. 41). Further, behavior that makes manifest the intention to make something manifest is *ostensive behavior*, or just *ostension* (p. 49). Finally, to *communicate*

---

[22] Davis holds that even beliefs that I am not now consciously entertaining are occurring to me (2003, p. 326). For this reason he may not find the example concerning Borel sets compelling. However, such a view of occurrent belief seems untenable, implying as it does that all those beliefs I hold but am not now consciously entertaining are occurring to me. For more discussion of Davis see Green 2007a.

[23] I use this account of manifestness rather than the official one that Sperber and Wilson offer on page 39 of their 1995, because that official one does not square with the use to which they put it. Their official definition has it that a fact is manifest to an individual just in case that individual is capable of representing it mentally, and capable of accepting that representation as true or probably true (p. 39). This definition counts as manifest any fact that I am capable of thinking about and believing true or probably true, regardless of my evidence. Accordingly, if someone outside my range of vision is walking towards me, then on this definition, the fact that he is doing so is manifest to me. However, Sperber and Wilson make clear that this is not their view; rather, the fact in question becomes manifest only when someone or something draws my attention to the person walking towards me, or does something that makes that approaching person visible (p. 49). For this reason I shall allow Sperber and Wilson's gloss of the notion of manifestness to stand in as a definition.

intentionally by ostension is to produce a stimulus with the aim of informing an audience of something, and intending to make it mutually manifest to audience and communicator that the communicator has this informative intention (p. 61).

An ostensive-communicative intention, then, for Sperber and Wilson aims at producing a cognitive effect on an audience, and we have seen in this chapter that this aim is not a necessary condition for speaker meaning. However, an intention to make mutually manifest to audience and communicator one's communicative intention does not require this; in just requiring mutual manifestness it demands only that one's intention be there for public view, so to speak. In this way it is like our notion of overtness. Yet it is unlike out notion of overtness in that it treats cases as ostensive intentional communication that our account of speaker meaning would not. An example is the case, considered above, of the river rat. In that example, Friend deposits a river rat in a house that he knows is being inspected for possible purchase by Homebuyer, and he does so in such a way as to make sure that Homebuyer observes him. However, Friend makes it look as if he thinks he is acting unobserved. In such a case we do not have speaker meaning, and our account of speaker meaning in 3.4 respects this fact. However, Sperber and Wilson's account of ostensive communication does not do so. To see why, observe that Friend (F) produces a stimulus aiming to achieve a cognitive effect in Homebuyer (H), namely the belief that the house is rat-infested (R). Further, F intends to make it manifest to H that he has this informative intention. After all, F intends H to be aware that he is trying to get him to believe that R. In fact, F intends to make it *mutually* manifest that he is trying to get H to believe that R.

Here is why: Let $R_{F \to H}$ be the proposition that F is trying to get H to believe that R. We know that $R_{F \to H}$ is part of F's cognitive environment, since surely F is either aware or capable of becoming aware of what he is trying to do. In addition, by the description of the case we know that $R_{F \to H}$ is part of H's cognitive environment. So $R_{F \to H}$ is part of F and H's shared cognitive environment. To show that $R_{F \to H}$ is part of F and H's mutual cognitive environment, we need only show that it is manifest to H, and manifest to F, that $R_{F \to H}$ is part of F and H's shared cognitive environment. (I consider just the case of H; the same reasoning will carry over to R.) This is easily done. H can surely see that $R_{F \to H}$ is part of his own cognitive environment. Can H see that $R_{F \to H}$ is part of F's cognitive

environment? Well, H can tell that F is trying to get him to believe that R; and presumably H knows that most people are aware of what they are trying to do. Consequently, H should be in a position to figure out that F is aware $R_{F \to H}$. We see then that the Sperber and Wilson approach treats the river rat example as a case of ostensive intentional communication, while the account of speaker meaning offered in 3.4 does not do so. That is good reason for preferring the present account to that of Sperber and Wilson, in spite of the fact that these authors correctly identify the intention to make facts manifest as an important element of communication.

4
Meaningful Expression

In the last chapter we developed a new account of speaker meaning as overtly showing an object, fact, or one's commitment—including both the modality and content of that commitment. Equipped with a viable conception of speaker meaning, in this chapter we'll first explore its relation to self-expression. Overt self-expression, it turns out, is a form of speaker meaning. We will also further refine the idea that in expressing ourselves we show our thoughts, feelings, moods, and experiential states. That refinement will enable us to substantiate the common-sense idea adduced in Chapter 2 that in expressing ourselves we in some cases quite literally make our ''inner'' states perceptible. As a result, the much-disputed contrast between the ''theory theory'' and ''simulation theory'' of how we know the intentional states of others rests on a false dichotomy: in some cases we know other minds simply by perception. Another consequence of the approach developed here is a new characterization of a little-explored kind of implicature—that phenomenon in which what we speaker-mean goes beyond what we say—that is neither conversational nor conventional.

4.1. Self-expression and speaker meaning

Chapter 2 glossed self-expression as showing and signaling an introspectible state. Given our definition of speaker meaning in Chapter 3, this implies that one who not only shows an introspectible state, but also overtly shows that state, both expresses and speaker-means that state. As illustration, recall the last chapter's story of Herod, who presents Salome with St John's severed head on a charger. Herod, we remarked, overtly shows that St John is dead (factual construal), or St John's severed head (objectual construal), thereby either speaker-meaning that St John is dead, or speaker-meaning St John's severed head. This suggests that were Herod overtly to

show not a severed head but his own feeling of, for instance, desire for Salome, he would not only be expressing his desire for her; he would be speaker-meaning that desire (or that he desires her) as well.

Again, you saunter in late to an important meeting of which I am in charge. I overtly scowl at you, not only making my anger manifest but also making manifest my intention to manifest it. Here it seems intuitively clear that I am expressing my anger. In addition, it seems intuitively clear that I am either speaker-meaning my anger, or speaker-meaning that I am angry. (Consistent with our discussion of the matter in Chapter 3, we need not assume that there must be a fact of the matter as to which of these characterizations is correct.) Our account thus far bears out and provides some rational reconstruction of these intuitions.

Much self-expression (either what is expressed or the act of so expressing it) is not, of course, a form of speaker meaning; indeed, even when an agent speaker-means something, what she expresses, in so doing, need not be speaker-meant as well. My inadvertent scowl, as I privately read irritating news, expresses my anger, but in so scowling I need not also speaker-mean that I am angry. Further, in a speech act such as assertion I speaker-mean some proposition P, and, in a way elucidated in 3.5, do so assertorically. In such a speech act, if I am also sincere I express my belief that P; we now know that this is a matter of signaling and showing my belief, and the theory of assertion as a handicap helps to explain how in performing that speech act I show, rather than merely signal, that belief. However, that assertion expresses belief does not imply that in speaker-meaning the content that I assert, I also speaker-mean that I believe what I assert. I sincerely assert: It's raining in Duluth. I thereby speaker-mean that it's raining in Duluth, and express, and thereby show, my belief that it's raining in Duluth. I can do all this without speaker-meaning that I believe it's raining in Duluth. In this way, what I express differs both from what I say and from what I conversationally implicate. As we'll see in more detail in 4.4, conversational implicata are part of what is speaker-meant. Yet because what I express in a speech act need not be speaker-meant, not all that I express in a speech act is conversationally implicated.

Philosophers of language, ethicists, and others stand at the business end of a long tradition in which what is expressed stands opposed to what is meant. Ethical expressivists, for example, hold that in calling something 'good' one expresses approval of the referent rather than predicating a property of it.

Even those not sympathetic with ethical expressivism will be inured to a perspective in which what is expressed stands beyond the pale of what is meant.[1] I take it as uncontroversial that some cases of self-expression are not cases of speaker meaning, and that some cases of speaker meaning are not cases of self-expression. What may remain controversial is the claim that one can express and speaker-mean one and the same object or state of affairs.

Objectual speaker meaning often takes as its object something observable—a charging bull, the looming funnel cloud. By overtly making manifest such a thing, we speaker-mean it. Does it make any sense, however, to suppose that a state that can be expressed, and thus be a possible object of introspection, can also be observed? I will argue that such a supposition does make sense, and in the process I hope it will emerge that making observable what's within is in fact a pervasive feature of daily life.

4.2 Showing what's within, part ii

We have elucidated the notion of self-expression in terms of the phenomenon of showing, which may take the form of showing that something is so (showing-that), showing something in such a way as to make it perceptible (showing α), and showing how an experience or emotion feels (showing how). As we saw in 3.5, some sincere speech acts express thoughts or other states of mind or heart by showing *that* one is in such a state. Besides showing-that, we also noted that showing can occur by making an object perceptible: I show you my bruised arm or the collapsed bridge. Is this showing-α form of showing germane to self-expression as well? If so, can it also be the content of an act of speaker meaning? I will offer an affirmative answer to both questions.

As observed in Section 1 of Chapter 2, it is also an integral part of our social life that we take ourselves to be capable of observing the feelings, experiences, and moods of others. There we noted that Amy Tan's narrator sees a crack of doubt on her sister GaoLing's forehead, and that Turgenev

---

[1] Wollheim 1968 nicely articulates this tradition in his description of what he terms a "dichotomy between expression and what is indifferently called communication, description, assertion" (p. 228).

tells us that the reproach in Arina Vlasyevna's eyes is visible. The perception of feelings is not limited to vision: The quaver in Hermione's voice enables us to hear her trepidation, and the intensity of Lionel's grip enables us to feel his exuberance. Offhand, such descriptions don't seem metaphorical. Had we, like the hammerhead shark, electroreception, we would surely also have coined a verb for the way in which that sense modality detects such things as muscular and nervous activity in other creatures, and we might even use that word to refer to our ability thereby to observe their emotions, moods, or experiences.

Or so it seems. The common-sense idea that we perceive one another's emotions might just be incorrect, worthy of the contempt we now feel for such outmoded ideas as that a person's bad mood is due to excess bile. Accordingly, let's see if it's justified. First of all, in what ways might showing enable perception of what is shown? I could explain to someone what my family, home, and neighborhood are like. If that explanation is sufficiently clear and convincing, I could show my audience how I live. I have not, however, thereby enabled them to see, hear, or otherwise perceive any aspect of my way of living. By contrast, I could show someone my home by having them over and walking them through its rooms, the attached yard, and so on. This latter means of showing something requires that I make that thing perceptible by means of one or more sense modality. Doing so enables someone to see or, with other senses such as touch or hearing, otherwise perceive my house.[2] Recall that enabling someone to perceive is one thing; getting them to do so is another. As we saw in the example of Ernest and Matthew of Chapter 3, it's not true that in order to show someone the looming funnel cloud, they must perceive it. Perhaps they are looking at the wrong thing or simply not paying attention. In that case they are being shown something without perceiving it. In what follows, when we speak of showing something in a way that enables perception—hereafter just *perception-enabling showing*—we do not assume that anyone does in fact perceive it.

Perception-enabling showing comes in at least two forms, making possible two corresponding ways of perceiving a thing. I shall refer to these

---

[2] A similar point is made by Stampe: "In saying truly that something is so, one does not enable his hearer to perceive, neither to see nor to hear, the way things stand. That however is precisely the kind of thing that such a device as a phonograph record or a telescope or an electrocardiograph does do" (1975, p. 234).

forms of perception as *part-whole perception* and as *perceiving in*. Under the former rubric, consider that an observable object, event, or process A that is a characteristic component of object, event, or process B can make B perceptible. Someone who presents to me the surface of an apple from one angle has thereby shown me the apple even if I do not inspect its interior or its other side. The reason is that a sufficiently large portion of a side of an apple is, for normal human observers, not only itself perceptible but also a characteristic component of that apple. This amounts to the fact that under normal conditions, perception of part of an apple's surface is enough to justify me in inferring (if only unconsciously) the existence of the entire apple. The facing side of a berry will likewise enable the bird to perceive the berry. On the other hand, because the bird does not normally perceive such small areas, a square micron of that facing side will not justify the bird in drawing this conclusion. That square micron will be a component but not, for purposes of perception, a characteristic component of the berry relative to the bird in its normal ecological situation. Again, a scanning electron micrograph of a bit of apple skin does not show a normal human observer an apple, even if it is attached to an apple, because that bit of skin is not a characteristic component of the apple relative to the human being in its typical ecology. However, if that observer is an experienced electron microscopist, that micrograph might well enable her to see it (supposing that the entire fruit is there to be seen). On the other extreme, I don't perceive the Milky Way by looking at my foot, even though the latter is a component of the former.

I might see the apple by seeing its facing surface; it does not follow that I have seen the *whole* apple. I have not seen any such thing, but we should be on guard against a confusion between seeing the apple and seeing the *whole* apple. The two events are not the same, nor is it plausible that in order to see the apple one must see the whole thing. You've seen the moon but not, unless you're an astronaut, the whole thing (even when it's full).[3] Let us keep in mind as well that whether one thing is, for purposes

---

[3] But is not the moon identical with the whole moon? Given the way we normally use the expression, 'the whole NP', the answer is no. 'The whole NP' is usually used to refer to all or most of the relevant parts of NP's referent; what is relevant in this case is what is perceptible. Accordingly, in the present example 'the whole tree' refers to all or most of the tree's observable surface, whereas 'the tree' does not. Were the issue testing the tree for termites, 'the whole tree' would refer to most of the interior as well as exterior of the tree.

of perception, a characteristic component of another will always be relative to a perceiving organism and its typical ecology. Hence a thing will, for purposes of perception, only be a characteristic component of an object *relative to an organism O in ecological situation E.* In the absence of such a relativization, asking what are a thing's characteristic components would be like asking for the time without specifying the time zone. It will help to make later discussion clear if we regiment usage as follows: Let α be an object, event, or process that is perceptible. Then we may say that relative to an organism O and ecological situation E, a *characteristic component of α* is a part of α that, when perceived in E without any other part of α being perceived, enables O to perceive α.

Part-whole perception may be distinguished from perceiving-in. For instead of being shown a thing by sensing one of its characteristic components, we may instead be shown a thing A by sensing a distinct object B in such a way that we see (or in some other way sense) A in B. We can see a horse in a mirror, in a telescope, and in a photograph (Walton 1984, 1997). Further, although the notion of showing is primarily bound up with vision, non-visual items may be perceived in other media. The sound recording allows us to hear the gunshot; perhaps we can even imagine an olfactory "camera" that detects, and then reproduces, the smells of a sumptuous meal not now present. Our ability to see, hear, feel, taste, and so on A in B is sufficient, though not necessary, for B's showing A. (In the case of a painting of a fictional horse, we make as if to see a horse in the picture but do not actually see any horse; likewise, very roughly, for music that is anguished. We shall return to this topic in Chapter 7.)

We've elucidated two ways of perceiving a thing: perceiving one of its characteristic components, and perceiving it in a distinct medium that is appropriately related to it. Clarifying the intuitive idea that we can observe one another's emotions thus raises the question which of these ways of perceiving makes this form of perception possible. Because of the importance of facial communication to our species and those from whom we most likely evolved, I shall pay considerable, though not exclusive, attention to the face in pursuing this question both in this chapter and the next.

We express many things in our faces, but one of the most familiar things we express is emotion, in particular one or more of the so-called basic emotions—standard cases of anger, fear, surprise, disgust, happiness,

sadness, perhaps also contempt. These emotions are basic in the sense that they are pan-cultural in our species, and thus evidently have a robust biological basis. In addition, these emotions are rightly discussed under the rubric, "affect programs". Following Griffiths (1997, ch. 3), that is to say that the basic emotions are complex, coordinated, and automated. Their complexity consists in their comprising a number of features, among the most important of which are that for each of the basic emotions we find:

(1) characteristic facial signatures,
(2) musculoskeletal, hormonal, and endocrine patterns,
(3) Autonomic Nervous System patterns.

The coordinated nature of basic emotions consists simply in the above features tending to occur in ensembles. Finally, the automatic nature of these basic emotions consists in their tending to occur without our conscious intervention. In many cases we can suppress these emotions' manifestations, but in general they are things that happen to us rather than being things that we do.[4]

To illustrate this "affect program" perspective, observe that many of our emotions have characteristic inputs, mediating states, and behavioral outputs. For instance, anger has typical causes such as perceived threat, typical subjective mediating states, galvanic skin response, increased blood flow to the body parts used for combat, and characteristic behavioral outputs such as attack (Frijda 1986). Likewise for the other basic emotions—fear, surprise, happiness, disgust, and sadness. Each of these emotions has, furthermore, a characteristic facial signature that is found with significant regularity across cultures.[5] The foregoing is of course consistent with the possibility of, for instance, anger occurring in the absence of its behavioral manifestations, since a disposition to behave in a certain way might never be manifested in behavior. It is also compatible with the possibility of behavioral manifestation occurring in the absence of anger, as in cases of dissimulation. Finally, anger is often automatic in virtue of being something that happens to us, sometimes even against our will.

---

[4] Ekman, P., Levenson, R., and W. Friesen 1983, Griffiths 1997, Frijda 1986, Bargh and Chartrand 1999.

[5] This notion of a program is best construed as "open" rather than "closed" in the sense of Mayr 1974. Also, we will consider the pan-cultural nature of the basic emotions in more detail in Chapter 5.

I wish to suggest now that not only can we know of one another's emotions; in some cases we can also literally perceive them. In particular I will suggest that in some cases we perceive emotions by means of part–whole perception. Given our definition of a characteristic component, simply assuming that a facial signature of, say, anger, is a characteristic component thereof would beg the question at issue: a characteristic component can only be a component of something that is itself perceptible. Instead, the proposition that a facial signature of anger is a characteristic component of anger is a hypothesis that, as we have seen, is supported by our common-sense patterns of thought and description of one another. To clarify this hypothesis, let's note that perceiving something does not imply perceiving all of its parts: I perceive your arm without perceiving the mitochondria that it contains. There might even be parts of a perceptible object that I cannot perceive: I can perceive a galaxy without, even in principle, being able to perceive the black hole at its center. So too the claim that we can perceive emotions is compatible with the possibility of their containing components that are not perceptible.

The perceptibility of emotions is consistent with the possibility that they can only be perceived by creatures with adequate conceptual repertoires. Thus, the perceptibility of bonobo fear does not imply that all human beings with properly functioning visual and other sensory modalities can perceive fear on the face of a bonobo. Instead, it's more plausible that emotions can only be perceived by those with appropriate conceptual resources. Furthermore, when an emotion is perceived, that does not guarantee that the creature perceiving it will know how that emotion feels. While it's plausible that in most cases in which one creature perceives an emotion in another, various automatic processes will occur that enable the former to know how the latter feels, those processes are not mandated by the perception itself. For instance, and as we'll see in more detail in Chapter 7, whereas "neurotypicals" tend automatically to empathize with those whose emotions they are perceiving, autists do not. Rather, autists can perceive emotions in the faces of others without knowing how they are feeling.

The thesis that emotions are in some cases perceptible swims against a strong and venerable current of thought in philosophy and related cognitive sciences. At the headwaters of that current is an old anxiety about the possibility of knowing other minds. It is an anxiety that lies

behind current controversy about a proper account, and it goes somewhat as follows:

We cannot observe cognitive, affective or experiential states directly. All we really observe is behavior. That behavior might in favorable circumstances flow from cognitive, affective or experiential states, but if it does, we need an account of the mechanism by which we justifiably draw conclusions about those states on the basis of that behavior. Short of such an account, belief in mental states of anyone but oneself will emerge as a mere dogma.

Current debate accepts this challenge at face value and attempts to meet it head-on. One approach tries to do so by suggesting that we draw conclusions about psychological states on the basis of behavior in the way that a biologist might form hypotheses about the nature of the cell (*theory theory*); another approach tries to solve this problem by suggesting that I project myself into the shoes of the person whose mind I am trying to understand, and then ask what I would think (feel, etc.) were I in that situation (simulation theory).

Current debate and the anxiety from which it flows assume uncritically that psychological states cannot be perceived. However, our observations thus far about expression, emotion, and the nature of showing call that assumption into question. For one, the alleged imperceptibility of all psychological states is not in fact a bit of common sense; everyday discourse goes the other way. Further, the common-sense presumption that emotions are sometimes perceptible deserves to be upended only if it turns out to be either internally incoherent or inconsistent with established fact. Certainly no internal incoherence jumps out at us. Let's consider some facts with which it is alleged to conflict.

Emotions cause such things as facial expressions. Should we infer that those facial expressions cannot also be characteristic components of those emotions? Such an inference would be fallacious. After all, it is perfectly natural to say both that the storm system caused the rain falling on my street, and that the rain falling on my street is a component of the storm system. At least according to ordinary parlance, then, the fact that an emotion causes a facial expression is consistent with its comprising that expression. It is even consistent with Hume's dictum that cause and effect must be "distinct existences", since of course a thing and one of its characteristic components are distinct.

One might instead object along the following lines to this suggestion that we can perceive emotions: If one could perceive, say by seeing, an emotion of, for instance, fear, it should be possible to say what color or colors that fear is; or if not (because one can visually perceive things in black and white) it should at least be possible to say or in some other way manifest what that emotion looks like. Similarly, if it were possible to hear emotion, it should be possible to say what an emotion sounds like.

We are, however, able to answer this question. One good way to do so is to draw a picture of a terrified face. Another way is to make such a face yourself. These responses count as perfectly adequate answers to the question, ''What did Peter's fear look like?''[6] Analogous remarks would be appropriate for the suggestion that we can hear emotions, and so on for other of our sense modalities.

Another source of resistance to the thesis that some emotions are perceptible might focus on their qualitative dimension. We've acknowledged that many emotions, including the basic ones, often characteristically have a qualitative feel: there is a certain way that rage feels to the person undergoing it; likewise for disgust, and perhaps also for happiness and sadness. Furthermore, how these emotions feel is not something that a third party can perceive with her senses: it is far from clear how I might go about observing the qualitative character of your emotion.

We have already found that a thing can be perceived even when one of its components cannot be: our (perceptible) galaxy contains a black hole. Instead, the imperceptibility to a third party of the qualitative dimension of an emotion is only a bar to the perceptibility of emotions if emotions are essentially or centrally qualitative, and all other components we have ascribed to them (behavioral tendencies, facial expressions, physiological

---

[6] One might also object to the thesis of the perceptibility of emotions along the following lines: Things that can be perceived visually reflect or emit light, things that can be heard reflect or emit sound, etc. However, it makes no sense to speak of anger as reflecting or emitting light, sound, or any other form of energy that might enable perception. Hence, the objection concludes, it is absurd to speak of perceiving emotions as well. In reply, I would urge that as speakers of one or more natural language we have no failsafe way of determining, simply by reflecting on our linguistic intuitions, whether a given expression makes sense. For all that those intuitions show, that expression might have an odd ring simply because it is very rarely used. Rorty's early work defending central state materialism against charges of incoherence makes precisely this point (Rorty 1965). Transposed to the current issue, the odd sound of such an expression as, ''Pete's rage reflects light,'' does not establish its incoherence. For all our intuitions can teach us, that odd sound might just be a result of our having not had much occasion thus far to speak this way.

responses, etc.) are at best consequences of those qualitative features. Even restricted to the basic emotions, however, it is difficult to see what could justify this privileging of one affective feature over others. From my own perspective, my first evidence about the presence in me of an emotion might be its qualitative dimension. This, however, only justifies a conclusion about the logical priority of the phenomenology of emotion if we assume that the order of knowledge mirrors the order of being. Rather than belabor the point, we need only note that this is a highly suspect assumption.[7]

Finally, it may be helpful to clarify why this argument on behalf of the perceptibility of emotions does not prove too much. In particular, an analogue of the argument would not support the absurd conclusion that we can perceive such cognitive states as belief. The reason is that belief has no facial, or any other behavioral, signature: there is no characteristic way that a person looks or otherwise behaves when she believes something. Instead, belief produces action via the medium of desire; and the variety of action thus produced is so wide as to be without any particular signature.

We have now made a case for the hypothesis that we can perceive emotions by means of part-whole perception. Does this hypothesis apply only to the pan-cultural expressions of basic emotions? In fact it does not. The perspective broached here is compatible with the existence of both individual and cultural variation in the display of emotion and other aspects of what is within. As we shall see in Chapter 6, we can make sense of the idea of perceiving an organism's emotion or other ''internal'' state without requiring that its manifestation of that state conform to a pattern across its species. For instance, even in convention-bound cases of expression it is possible literally to see emotion in a person's face.[8]

---

[7] This defense of the perceptibility of certain emotions is not meant to marginalize ways of expressing ourselves other than those by which we make states of ourselves perceptible. We have already highlighted the way in which speech acts enable us to show that we are in a certain psychological state. Further, as we will see in Chapter 7, one can also express oneself by showing how one's emotion feels.

[8] This point is made by Hampshire 1971, p. 145. Also, as observed in Wollheim 1968, we can imagine a situation in which we manifest anger by means other than the facial configuration we currently associate with scowling. This would be a world in which, say, people normally manifest anger by means of a face that in the actual world we associate with smiling. This possibility is in no tension with our account of our ability to see anger in a scowl. Were that possibility realized, we would then see anger in (what we now call) a smile, and this would be due in part to the fact that smiles are in this situation characteristic components of anger.

This way of approaching matters also points toward other, less biologically grounded ways of showing emotion that are nevertheless not bound by convention. Developing an earlier example, suppose Elaine is conducting a meeting to which her colleague Juan arrives late. She shapes her hand into a gun and pretends to shoot him. This is a way of showing, and indeed expressing, annoyance and we are now in a position to see how it is achieved. It is manifest that Elaine is making as if to dispatch Juan, and as discussed in Chapter 2 we often make as if to do things for the partial satisfaction that this can give. It is thus natural to see Elaine as vicariously satisfying her urge to dispatch Juan, and her action shows that urge. But that urge is a characteristic component of her annoyance, a relatively mild species of anger, so Juan and his colleagues can see Elaine's annoyance in her pretence. We return to examples of this sort in Chapter 6.

As we also observed in Chapter 2, some phenomena show one's intentional state without expressing it. Blushing shows, perhaps also betrays, one's embarrassment, but one does not express embarrassment by blushing because it does not appear to be a signal. We also remarked in that chapter that two species of expression correspond to two ways in which a manifestation of a thought or emotion might be intentional. On the one hand, if a person scratches out the eyes in a photograph of a rival, she is expressing hatred (and quite possibly exorcizing some of that hatred as well). Assuming that she is not attempting to practice voodoo, she is not intending to communicate and her action is not a case of speaker meaning. In contrast with the non-overt variety of expression, we also express ourselves for the sake of manifesting to others our cognitive or affective state. One line of thought in Grice, however, would seem to challenge the possibility of this latter phenomenon, so I'll now consider it.

4.3 Expression and automaticity

We saw in Chapter 2 that some self-expressions are voluntary acts, that is, things we either do intentionally or allow to happen. There we also claimed, and since then we have been developing the thought, that in self-expression we show a psychological state, sometimes in such a way as

to make that state perceptible. However, our gloss of the "characteristic component" species of showing involves manifesting a state that has been described as part of an "affect program". In this case, it might seem strange that components of programs to which we are subject should also be under our voluntary control.

No doubt certain manifestations of happiness, anger, surprise, and so on can occur whether we want them to or not. (In the case, for instance, of smiling, we have evidence that the "Duchenne Smile" involves contraction of facial muscles that cannot be contracted at will; more on this in Chapter 5.) However, the definition of a program in the sense relevant here is compatible with the actions comprising it being voluntary. Consider an analogous functional view for propositional attitudes such as belief. Action taken in light of a belief is action nonetheless, and likewise for desire. Furthermore, in very few cases are we so angry (to take just one case) that we literally lose control of ourselves. Rather, more commonly our emotions provide us with impulses of varying intensity on which we may or may not choose, or allow ourselves, to act.

There is no contradiction, I suggest, between our behavior being part of a program and its being voluntary. Further, reflection on daily social life suggests that a great deal of our expressive behavior contains an element of the voluntary. Many of our smiles we make, many of our scowls we produce at will. We put on a look of surprise when someone says something out of place (recall the "ironically freaked-out" face of Section 1.2); we make a pouty face when someone we care for hurts our feelings. Further, we often exaggerate the facial configurations that do happen to us, as for instance when I deepen my frown in order to underline the gravity of someone's infraction.

Some authors seem tacitly to suppose that to the extent that our expressive behavior is voluntary, to that extent we must be dissimulating. The possibility of voluntary expression certainly leaves open a space for mendacity. If I see someone I am only slightly glad to encounter, and to curry their favor I put on an ear-to-ear smile, I am dissimulating. At the same time, however, many of our emotions may be felt deeply without having any immediate involuntary consequences. I might for instance be thoroughly irritated with a friend's unpunctuality without becoming enraged. When I give my best Clint Eastwood scowl as she arrives late for

a meeting yet again, I am giving vent to that irritation without its forcing itself out of me. On the other hand one could not accuse me of mendacity in scowling as I do.

I am contending that with no departure from sincerity we can express ourselves voluntarily. One line of reasoning in Grice suggests otherwise, and I shall pause to examine it. Grice seems to hold that one producing a frown intentionally can only reasonably intend it to provide an audience good reason to infer that the frowner is displeased if she has reflexive communicative intentions. He writes,

If I frown spontaneously, in the ordinary course of events, someone looking at me may well treat the frown as a natural sign of displeasure. But if I frown deliberately (to convey my displeasure), an onlooker may be expected, provided he recognizes my intention, *still* to conclude that I am displeased.… [T]hough in general a deliberate frown may have the same effect (with respect to producing belief in my displeasure) as a spontaneous frown, it can be expected to have the same effect only *provided* the audience takes it as intended to convey displeasure. That is, if we take away the recognition of intention, leaving the other circumstances (including the recognition of the frown as deliberate), the belief-producing tendency of the frown must be regarded as being impaired or destroyed. (1957, p. 219)

This passage comes in the course of an argument for the necessity of reflexive communicative intentions for speaker meaning. For this reason, Grice seems to be construing 'intended to convey displeasure' as 'intended to produce in the audience belief in the frowner's displeasure by means of recognition of this intention'. The claim, then, is that the deliberate frown can only be expected to have a belief-producing tendency if it is produced with the intention of serving as a testimonial of the frowner's displeasure—as an unspoken statement. I do not know whether Grice considered another reading of 'intended to convey displeasure', but one is available. Since 'conveying displeasure' could also be read as 'manifesting displeasure', an intention to convey displeasure could also be taken as an intention to manifest it rather than testify to it. Whether or not Grice considered this question, we do well to ask whether one intending to manifest displeasure can reasonably expect her frown to have a belief-producing tendency.

Behind Grice's reasoning seems to be the assumption that if A construes U's frown as intentional but not involving reflexive intentions, A will be

led to conclude that U's intention is instead to deceive, or at least will have strong evidence in favor of this conclusion.[9] This, however, is dubious. As we have just seen, displeasure, like many other emotions and moods, has natural manifestations that are not so powerful that they force themselves out of us involuntarily. In many cases such natural manifestations can be inhibited. Further, when those manifestations are not inhibited but, (a) at the time they are manifested, could have been, and (b) we refrain from inhibiting them for a reason, then they merit treatment as intentional. Suppose now that U feels anger in response to something that A is doing, but not such anger that the feeling manifests itself involuntarily, that is, against his will. Nevertheless, in order to get A to refrain from the offending action, U allows himself to do what his anger disposes him to do, and he scowls. Suppose further that A recognizes both the scowl and the fact that it has been produced voluntarily.[10] Knowledge of the scowl's voluntary nature need not undermine the inference that A might make as to U's affective state; knowing that anger tends to produce both voluntary and involuntary scowls, U could justifiably form the belief that the scowl is A's way of intentionally manifesting his anger. What is more, if A has doubts about the authenticity of U's facial expression, it is not clear how they will be assuaged by belief that it is produced with the intention that A come to believe that U is angry by means of recognition of that intention.[11]

---

[9] That Grice is assuming this is further suggested by his reasoning in his later piece, 'Meaning revisited'. In discussing the difference between an involuntary expression of pain and a voluntary expression thereof, Grice writes,

In stage two not only does creature X produce this behavior voluntarily instead of nonvoluntarily, as in the primitive state [stage one], but we also assume that it is recognized by another creature Y, involved with X in some transaction, as being the voluntary production of a certain form of behavior the nonvoluntary production of which evidences, say, pain. That is, creature X is now supposed not only to simulate pain-behavior, but also to be recognized as simulating pain-behavior. The import of the recognition by Y that the production is voluntary undermines, of course, any tendency on the part of Y to come to the conclusion that creature X is in pain. (1989, p. 293)

I suggest we question Grice's assumption here that pain behavior that is voluntary must be a simulation of pain behavior not caused by genuine pain. Once that assumption is questioned it will become quite unclear why Y's recognition of X's production as voluntary will undermine any tendency to produce the belief that X is in pain.

[10] We have strong evidence that people are able to detect when certain facial expressions are produced voluntarily. For discussion and meta-analysis of other studies see DePaulo 1992.

[11] Social psychology, particularly that influenced by Goffman, has come to appreciate the importance for smooth social interaction of deliberate manipulation of expressive behavior. For further discussion see DePaulo 1992 and DePaulo and Friedman 1998. We note here also that, as argued in Frank *et al* 1993, there may well be physiological markers distinguishing involuntary smiles from those produced

We said in the last section that we overtly express ourselves for the sake of manifesting our cognitive, affective, or experiential states. It is now clear that U, grasping the line of reasoning that we have just articulated, may know that an audience A can discern a voluntary performance of his as just such a manifestation. As a result, U could do or allow something with the intention of manifesting her cognitive or affective state, and if her intention is fulfilled she has expressed that state.

Expressing a state of heart or mind is not, of course, necessary for speaker meaning. One can, for instance, illocutionarily speaker-mean something without being in the state one's words purport to express. One making an assertion shows her commitment, but also purports to express a belief; the liar thus shows her commitment to a proposition that she does not in fact believe. What is more, there seem to be many speech acts that neither express nor purport to express any state of thought or feeling. (Consider, for instance, appointing or excommunicating.) As we have seen, another element that is not necessary for speaker meaning is an intention to influence anyone's beliefs. The relations among the notions of speaker meaning, illocution, and self-expression are set out in Figure 4.1.
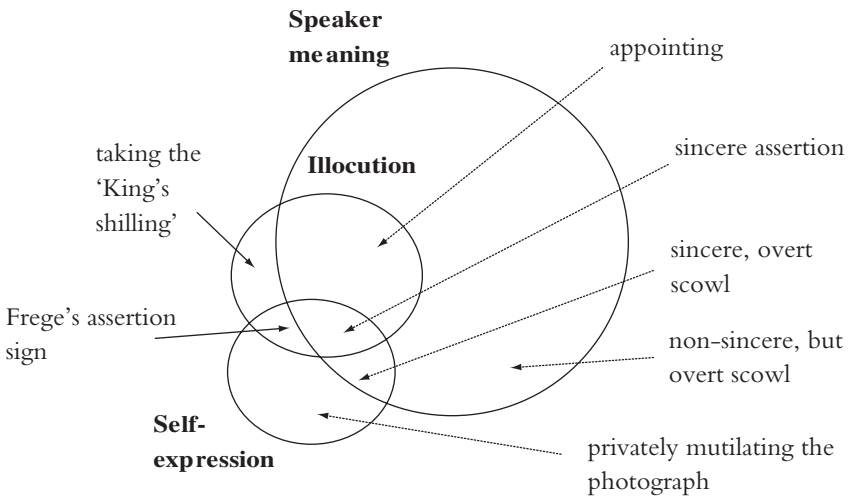

Figure 4.1: Relations among speaker meaning, illocutions, and self-expression

in part by voluntary means. This does not imply, however, that detection of smiles in the latter category must be detection of dissimulation. We pursue this point further in Chapter 5.

The following remarks are intended to elucidate the annotations on Figure 4.1:

appointing. Appointing is a clear example of a speech act. It meets the well known criterion of being an act that can be performed by saying one is performing it. "I appoint…" uttered under the right conditions is sufficient for appointing. At the same time, one who appoints someone to a post or position must do something by virtue of which they mean that the appointee is to be thus appointed. The person doing the appointing must thereby produce a case of speaker meaning. Hence appointing is both an illocution and a case of speaker meaning. However, there is no attitude that is characteristically expressed by an appointment. This is suggested by the fact that it is hard to know what could be meant by one who asks of a speaker who, by her words, has succeeded in appointing someone to a certain post, whether she is sincere.

sincere assertion. An assertion is an illocution, passing inter alia the test that one can assert that P by saying that one is doing so. It is also a case of speaker meaning, since one cannot assert that P without meaning that P. Finally, if the assertion is sincere, then the assertion expresses the speaker's belief that P.

sincere, overt scowl. Producing an overt scowl is a way of meaning that one is angry. That is evidently the point of making the scowl overt, rather than just publicly accessible. Further, if it is sincere, that must be because the person scowling is angry. It would also seem that the scowl is an expression of the person's anger; in that case the overt scowl is both a case of speaker meaning and of self-expression. However, there is no speech act of scowling. (One cannot scowl by saying, 'I hereby scowl.')

non-sincere, but overt scowl. This case differs from the former only by virtue of not being one of self-expression. One who scowls without being sincere is not expressing *her* anger.

privately mutilating the photograph. When Jane privately mutilates the photograph of Judy, with no intention that Judy or anyone else see the fruits of her labors, she expresses her anger at Judy, and so this is a case of self-expression. However, it is neither a case of speaker meaning nor an illocution.

Frege's assertion sign. Frege held that one using his logically correct language, the *Begriffschrift*, must mark the development of his reasoning with the ''assertion sign''. That sign is used to indicate one's acknowledgment of the truth of the proposition that it prefixes. In this way it has an expressive dimension: one who uses it shows and signals, and thus expresses his acknowledgment of the truth of a proposition. The use of the assertion sign is also an illocution: one who uses it asserts the proposition that follows it. By contrast, the use of the assertion sign is not necessarily a case of speaker meaning. One can use it without intending thereby to display overtly one's belief in the proposition it prefixes.[12]

taking the ''King's shilling''. During a certain period in British history, it was possible to enlist people for the navy by causing them to accept money from a representative of the navy. By law there was no further requirement that one who enlists be aware of enlisting in the navy or intend to enlist in the navy; apparently it was not even necessary to know that one had accepted the money. Thus for instance if a peasant is drinking in a pub with (what is unbeknownst to him) a recruiter for the navy, and while the peasant is looking away the recruiter slips a shilling into his ale, then upon drinking that ale the peasant has enlisted in His Majesty's Navy. Since one can also enlist by volunteering, and one can do this by uttering the words, ''I volunteer'' under the right conditions, enlisting is an illocution. It follows that enlisting is an illocution that one can perform unawares, and thus is an illocution that can be performed without also being a case of speaker meaning.[13]

4.4 Expression and implicature

We have seen that while speech acts such as assertion have an expressive dimension, expression is not itself a species of assertion. Rather, self-expression also lies beyond the pale of what is said.[14] In expressing disgust by making a disgusted face, even overtly in such a way that we

---

[12] Green 2001 defends this view of Frege's assertion sign in further detail.

[13] Green 1997 considers in further detail the possibility of such ''strict liability'' speech acts.

[14] Saying and asserting are to be distinguished by at least this much: one can say that P without asserting it, as for instance when one utters a sentence in the course of making a supposition for the sake of argument.

(speaker-) mean that we are disgusted, we do not thereby say that we are disgusted any more than we assert that we are. This raises the question how self-expression is related to another phenomenon in which what we mean goes beyond what we say, namely implicature. On its orthodox treatment, when what is meant is not part of what is said, that can be due either to the conventional meaning of the expression used, if such there be, or it can be due to non-conventional features of the expression uttered together with the context in which it is uttered. The former is conventional implicature, while the latter is non-conventional implicature. Students of implicature often identify non-conventional implicature with conversational implicature, and conventional and conversational implicature have been the central subjects of research on implicature. What is not often mentioned[15] is that Grice's framework leaves room for a form of implicature that is neither conventional nor conversational. A diagram might help (see Figure 4.2).

Speaker meaning

what is said   what is implicated

conventionally   non-conventionally

conversationally   non-conversationally
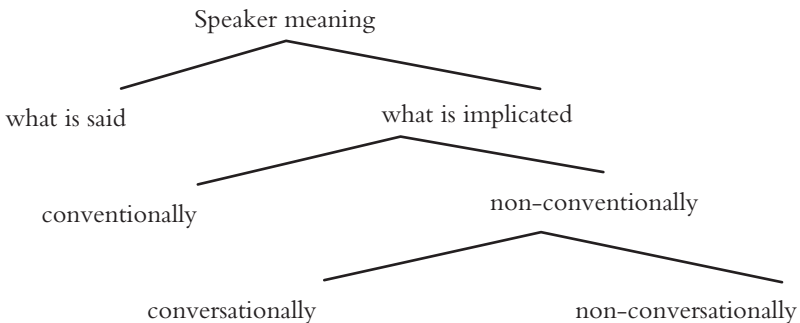
Figure 4.2: Forms of speakes meaning

When I overtly scowl, then as we have seen, my act may be both an expression of anger and a case of speaker meaning. We have also observed that in so scowling I need not be saying or asserting that I am angry. In addition, I can manifest my anger without exploiting conversational norms in the way that conversational implicata do. If I am less informative than I might have been expected to be, or apparently irrelevant, I may convey something distinct from what I say by virtue of this apparent violation of conversational norms. By contrast, my overt manifestation of anger does not depend upon any apparent or genuine violation of conversational norms.

---

[15] An exception is Green 2002.

As a result, my speaker-meaning that I am angry in scowling is not a species of conversational implicature. Nor is it due to any conventions governing scowls. (As we've seen briefly already, and as will be developed in further detail in Chapter 5, some emotions can be displayed in the face without the aid of conventions.) As a result it is also not a form of conventional implicature. If Grice's taxonomy of forms of speaker meaning is exhaustive, then, it will follow that the speaker-meaning variety of self-expression is a form of non-conventional, non-conversational implicature.

That a phenomenon falls into a category does not immediately help to shed light on it. Whether we have an explanation of how a conversational implicatum is achieved depends on how well our theory of conversation works; likewise, whether we have an explanation of how a conventional implicatum is achieved depends on how well our theory of conventions works. By placing an overt facial expression of an emotion into the category of non-conversational, non-conventional implicature, we illuminate it by invoking the notion of showing: I speaker-mean what I do in such a case by overtly showing my state of mind (heart, experience). So far we have considered how this is done either by showing-that or by showing-α. How far this explanation takes us depends, of course, on the adequacy of our elucidation of the notion, or notions, of showing.

That cases such as the scowl we have just been considering are forms of implicature, is consistent with other more widely discussed forms of implicature being instances of self-expression also. For instance, conventional implicature is commonly described as an instance in which a speaker means something without saying it, while her ability to mean this is due to the meaning of her words. Thus in using 'but' in a sentence such as 'Mary was poor but honest' a speaker will suggest that there is, or is thought to be, some tension between poverty and honesty. This is due, further, to the meaning of that conjunction. It's no accident that words like 'but', 'therefore', 'nevertheless', and many expletives, which are paradigms of conventional implicature, are often also described as expressive devices. In Section 6.6 we will explore a case of one verbal device that may be construed as designed for conventionalized self-expression.

We are now also in a position to see that conversational implicature, when sincere, is a form of self-expression. In support of this contention, observe that a conversational implicatum is typically not "marked" as having one or another illocutionary force: Conversational implicata are

not all cases of indirect speech act. To see why, observe that it does not seem accurate to say that when we implicate that we cannot be any more informative than we are being in answer to a request for information, we are asserting without saying that we cannot be any more informative. After all, if we were asserting this, but knew that we could be more informative, we would be lying. Yet an insincere implicatum is not, insofar, a lie. In saying, ''Somewhere in Kentucky'', U may fail to be as informative as he can have been expected to be in answer to the direct question, ''Where is John?''. So far it is not clear that U means that he can be no more informative, since U may have been reticent for any of a variety of reasons. However, if U is attempting to be a cooperative interlocutor, U's reticence can be justified if he believes that he can be no more informative without infringing a conversational maxim such as Quality (''Say only that for which you have adequate evidence.''). Given the presumption that U is attempting to be cooperative, U may reasonably intend his reticence to show his belief that he can provide no more information than he has given. He might also intend to make A aware of his inability by means of recognition of this very intention, but this further intention is not mandatory. It seems sufficient for U to mean that he can be no more informative that U intend his reticence to show his belief that he can be no more informative. Rather than overtly showing a bandaged leg, U overtly manifests the infirmity of his information; in both cases something is speaker-meant, and in the latter case something is both expressed and implicated.

This last case of overtly being less informative than one is called upon to be involves a speaker knowingly violating a conversational maxim (Quantity) due to his inability to fulfill it without violating a more stringent one (Quality). Other cases of implicature involve not merely violating but flouting a maxim, and we do well to consider the form of speaker meaning that they characteristically involve. What a person says might contain less information than required, but if they also may be expected to be fully informed about the situation they are discussing, and can be assumed to be cooperative, then their reticence cannot be due to a clash of conversational maxims. A more likely hypothesis is that their reticence is aimed at avoiding a clash with another maxim, say of etiquette (they fear offending someone) or of prudence (they fear liability for a lawsuit). Only those considered authoritative about their subject matter damn with faint praise. Thus one writing a letter of recommendation, or one giving a

verdict on a dish they are sampling ("Well, it was nicely presented on the plate"), most likely knows all she needs to about her subject matter. In such cases speakers implicate that the object they are assessing falls on the low end of the scale salient in that conversation. This way of describing what they are doing, however, is undecided between construing them as overtly manifesting the belief that the object in question falls on the low end of the salient scale, or as asserting (without saying) this. Since, in attributing a simpler set of communicative intentions, the former characterization is more parsimonious while being equally explanatory, we do well to accept it. Further, as with the example of John's location in Kentucky, viewing the implicatum as silently asserted would predict that if the speaker did not believe what he implicated, he would be a liar. However, when I, for instance, damn with faint praise while knowing that I could be more complimentary, I am worthy of abuse but still not guilty of lying.

In both cases we have considered, it is of course possible for the speaker conveying the implicatum in question to harbor reflexive communicative intentions. Nothing prevents a speaker having more complex communicative intentions than what is necessary for implicature. Our claim is not, then, that telling and implicating are incompatible. Rather, it is the weaker contention that a speaker may convey a conversational implicatum merely by using her presumed adherence to conversational norms, together with her failure to adhere to all conversational requirements, overtly to manifest her state of mind. She might also intend to make a claim on her addressee's views, but doing so would go beyond what is strictly necessary for either speaker meaning or implicature.

It might be suggested that we can trim down the attribution of communicative intentions even further than I have been advocating. Gauker 2001 proposes just such a strategy, arguing that alleged standard cases of conversational implicature do not require description in terms of implicature at all. Rather, he suggests, communication can succeed by means of a speaker putting forward a propositional content and the addressee inferring a conclusion on the basis of that content together with situational factors. (This inference he calls "situated inference".) For instance, one looking for her friend might ask, "Where's Mary?" and receive the reply, "I've heard gardening noises out back." From the proposition that gardening noises are coming from the back yard, together with other background knowledge about the current situation, the addressee may with little difficulty infer

that Mary is probably in or near the backyard. According to Gauker, no reference to the speaker's intentional states, to say nothing of intentions to produce beliefs in the addressee by means of recognition of those very intentions, need be made by that addressee. Rather, he writes, ''in all of the sorts of cases where Grice finds a conversational implicature, we might instead suppose only that the hearer draws an inference from what the speaker literally says and the external situation.'' (2001, p. 164.)

In certain cases communication can succeed by means of a speaker's uttering a sentence and an addressee's drawing an inference from the content of that sentence together with facts about the situation in which it is uttered. However, in many cases it is doubtful that these two factors will be enough for the speaker to convey the wanted message. Sometimes the addressee will need to consider hypotheses about why the speaker is asserting one content rather than some other, and in so doing she will need to attribute communicative intentions to the speaker. For instance, A might say, ''The dog has done something on the carpet.'' It is far from clear how this content, together with situational factors, will allow the addressee B to infer that the dog has relieved itself on the carpet, rather than any of the many other things that it might have done on the carpet, such as scratching itself, panting, eating a bone, or barking. For B to discern what A is trying to convey, B will have to consider why A asserts an otherwise unremarkable fact. Barking, scratching, panting, and bone-eating on carpets are generally dog actions not requiring responses from their owners, whereas voiding on carpets is. This fact will help B determine why A has made the remark that she has, namely to manifest her belief that the carpet has been soiled, this attribution to A being required for communication to succeed in this case. Grice sometimes speaks of calculating conversational implicature as requiring an inference on the part of the addressee not from what a speaker has said, but from her saying of it, and Gauker's proposal fails to respect this crucial element.

4.5 Alternative accounts of self-expression

As we did in Chapter 3 after developing an account of speaker meaning, I here pause to consider rivals to my account of self-expression. In each

of the three cases below I will argue that the alternative account has some merit but is nevertheless inferior to the account being developed here.

Bach and Harnish on self-expression

In their influential 1979, Bach and Harnish place great weight on the notion of expressing. In particular these authors taxonomize illocutionary acts in terms of the type of attitude they express: constatives express the speaker's belief, directives express the speaker's attitude toward a prospective action, and so on (1979, p. 41). Their account of expression is as follows:

> In uttering x, S expresses attitude A iff S utters x with the intention that an addressee H by means of recognizing this (entire) intention, take the utterance as reason to think that S has A. (1979, p. 15)

We have already seen that one can express an attitude without speaker-meaning that attitude. For instance, one can sincerely assert that P, thereby expressing one's belief that P, without thereby speaker-meaning that one believes that P. However, Bach and Harnish assume that what one expresses one also speaker-means. This is incorrect. Further, we learned in Chapter 2 that even if these authors are correct in holding that an expressed attitude must also be speaker-meant, they would still be incorrect to construe speaker meaning as requiring reflexive communicative intentions.

Bach and Harnish's elucidation also fails to distinguish between producing an utterance with the intention that it be taken as evidence of one's attitude A, and producing it with the intention that it be taken as showing that one has attitude A. Further, this definition takes no account of the difference between showing that one is in an intentional state A, and showing that state in such a way as to make it perceptible.

Davis on self-expression

Davis 2003 develops a notion of expression intended as a crucial element in his refurbishing of the ideational theory of meaning. On this approach words have meaning by virtue of expressing ideas; sentences by virtue of expressing thoughts. Davis holds that self-expression is essentially an intentional phenomenon, and remarks as well that because facial expressions of emotion are typically involuntary manifestations of emotion (§3.1), he

will set them aside as not directly germane to the concept he wishes to elucidate. We have seen two reasons for thinking this an unfortunate restriction. First of all, as argued in Chapter 1 and developed further in this chapter, our intuitive notion of self-expression includes things that we allow as well as things that we do. Depending upon one's use of jargon, things that we allow may or may not be intentional, but they do yield self-expressions if we allow them to occur for the sake of showing how we think or feel. Furthermore, as we shall see in Chapter 5, the alleged involuntariness of facial expression disguises two ideas: First, the "can't do it at will" sense, and, secondly, the "can't help it" sense. Many facial expressions that we are inclined to call involuntary are so in the first but not the second sense, but behaviors that are involuntary in this first sense are nevertheless things that we allow, and often things that we allow for a reason. This is one reason why we pay close attention to facial expression in the present work.

Next, Davis uses 'expression' to refer to cases in which people purport to express thoughts or feelings they lack. On this usage insincere expressions are possible. As a result it won't do for Davis to elucidate expression, as we do, as showing how things are within, since 'show' is a success verb while on his usage 'express' is not. Instead Davis uses the notion of indication, in which A's being an indication of B is evidence, but not conclusive evidence, that B holds. Further he holds that the indicator A must be publicly observable, with the intended consequence that self-expression requires producing something observable by others. This is in agreement with our own approach, as is his observation that what is observable need not in fact be observed.

Next, Davis points out that an intentional act that happens as a matter of fact to indicate a state of thought or feeling is not yet a self-expression. (Recall our example of Portia, whose lighting a cigarette indicates that she is nervous, but whose action is not an expression of nervousness; see also Kemmerling 2002.) Instead the act must be done with a (not necessarily the) purpose of providing an indication of a state of thought or feeling. Not only that, but the action must be intended as itself such an indication rather than as a means to providing such an indication. Thus one who withdraws cash from her bank to buy flowers for her beloved performs an intentional act for the purpose of providing an indication of her affection, but her cash withdrawal is not an expression of affection. Davis shares

with our approach as well the requirement that expression, like speaker meaning, must be overt. What is to be ruled out are covert simulations of unintentional indications of an attitude.

Davis holds in addition that one can only express occurrent states. For instance for the case of belief he holds that one can only express the belief that P by also expressing the occurrent thought that P. Then letting Ψ be a term for an introspectively applicable concept (a concept one can ascribe to oneself on the basis of introspection), Davis gives us the following general definition:

> S expresses Ψ iff S performs an observable action as an indication of occurrent Ψ without thereby covertly simulating an unintentional indication of Ψ.

As Davis remarks, the observable action in question need not in fact succeed in being an indication of the agent's Ψ; since, as he holds, expressions are actions, in the nature of the case they might fail. However, while in some cases when they fail we might say that the agent failed to express herself fully or adequately; in other cases an agent's actions might fail to express her thoughts or feelings *simpliciter*. Thus recall the example of a painter who feels desolate but winds up producing a picture of serenity, due to a wrong choice of color scheme. It is not as if, were we to back up far enough from what happens on the canvas to include in our view her central nervous system, we could be assured of finding acts that do in fact express her desolation. She does not express her desolation at all in spite of trying to. Nevertheless Davis's account treats her as doing so, and this is incorrect. Further, on our approach the idea of performing an action that is an unsuccessful attempt to express oneself makes clear sense; on Davis's approach the only sense that can be made of a failed attempt at expression is that of attempting but—say due to momentary paralysis—failing to perform an action intended to achieve self-expression.

Davis's conception of self-expression is also narrower than ours because it rules out non-overt cases. When Jane scratches out the eyes in a photo of her former friend she performs an observable action but not for the purpose of providing an indication of an occurrent mental state. Rather, she is making as if to act on her rage in order vicariously to gratify it. Her action nevertheless seems to be a clear case of expressing rage, albeit not

for communicative purposes. Davis's gloss of the notion of expression does not allow for a case of this kind.[16]

Kemmerling on expressing an intentional state

Andreas Kemmerling (2002) offers the following gloss on the notion of self-expression:

> In doing x, S expresses the Ψ that p iff in doing x, S does something which is (incontestably) cp-analytical evidence of his Ψ-ing that p.

For Kemmerling, something is cp-analytical just in case it is false as it stands but can be turned into a conceptual truth by adding normality qualifications. Thus even if not all assertions are evidence of belief, a normal assertion arguably is; and indeed it seems an analytical truth that normal assertions express belief. Kemmerling's account thus rules out cases of expression that are due to contingent biological regularities. As we shall see in Chapter 5, specific facial expressions are associated with certain emotions, but even when conditions are normal that fact is not a conceptual truth. We also know that while many such facial expressions are not voluntarily produced in the "can do it at will" sense, they are voluntarily produced in the "can help it" sense. As a result a suppressible Duchenne smile (which is such a case) will not count as an expression of pleasure on Kemmerling's approach.[17]

Our discussion thus far points to a number of further questions to be discussed in the next chapter. First of all, a major school of thought—the Neurocultural View—concerned with the psychology and neuroscience underlying facial expression holds those expressions to be involuntary. We have taken an opposed position, and it will be fruitful to examine the basis for this alternative view. Further, an emerging view—the Behavioral Ecology View of facial expression—opposes the aforementioned school of psychological thought, denying that facial expressions are ever manifestations of emotion, and holds instead that they are always, and only strategically guided messages. The Behavioral Ecology View also rejects the intelligibility of questions about the sincerity of facial expression. We in contrast have taken self-expression to be essentially sincere, at least in cases

---

[16] I critically examine Davis's position in more detail in Green 2007a.
[17] See also Tsohatzidis 2004 for a distinct objection to Kemmerling's account.

in which one expresses one's thought or feeling, and so once again do well to examine the merits of an opposing position. It will emerge that these two schools of psychological thought can with some modification be shown compatible, and that the key to seeing this possibility is accepting our view of some forms of self-expression as intentional and overt manifestations of what is within.

5
Facial Expression

We have been developing and refining a view of self-expression as signaling and showing psychological states. In the last chapter we learned that self-expressions that are also speaker meanings are done or allowed with an intention of overtly manifesting what's within; in other cases of self-expression we find doings, allowings, or other behaviors that as a matter of fact manifest what's within, but are not performed with an intention of doing so. One can, indeed, intend overtly to manifest something without also intending to have any effect on the beliefs or other psychological states of others. Further, sometimes what we manifest in the course of expressing ourselves we also make observable. Introspectible states are, therefore, also in certain cases observable. Further, introspectible states, when the object of (objectual) speaker meaning, also provide clear cases of non-conventional, non-conversational implicature; other species of implicature are also best conceptualized in terms of self-expression.

In this chapter we consider how our theory of emotions as literally perceptible in such places as the face holds up in light of two currently influential approaches to the experimental psychology of facial expression. We will see that our account can be clarified and refined with the aid of insights from each of these approaches; at the same time it can be used to overcome errors in each of them in such a way as to suggest a third approach to facial expression combining the best insights of the two current approaches.

Many higher primates have an impressive repertoire of facial expressions, and for human beings a paradigmatic locus of self-expression not requiring conventions is the face. Of the approximately forty human facial muscles, most are attached only on one side to bone, and on the other to facial skin or fascia. It has been suggested accordingly that these facial muscles are specialized for communication and expression (Rinn 1984). This suggestion in turn raises the question whether facial expression is pan specific, or is on

the other hand variable from one member of our species to another or from one culture to another. For the first two-thirds of the twentieth century, it was widely supposed that virtually nothing of interest is held in common among the practices of facial expression from one individual or culture to another. However, in the late sixties a group of researchers including Paul Ekman, Carroll Izard, and Wallace Friesen began to make a powerful case for universals governing facial expression for certain emotions. It then became natural to hypothesize that facial expression is governed in large part by our biological endowment and is, further, a product of natural selection. This spurred a renewal of interest in a comparatively neglected work of Charles Darwin, and to which we now turn. It will facilitate discussion to have a diagram of the main facial muscles, and I reproduce one such in Figure 5.1.

Occipitofrontalis

Levator labii
superioris

Zygomaticus
minor

Zygomaticus
major

Buccinator

Levator anguli oris

Depressor labii
inferioris

Temporalis

Procerus

Orbicularis oculi

Levator labii
superior alaeque nasi

Orbicularis oris

Masseter

Depressor anguli oris

Mentalis

Platysma
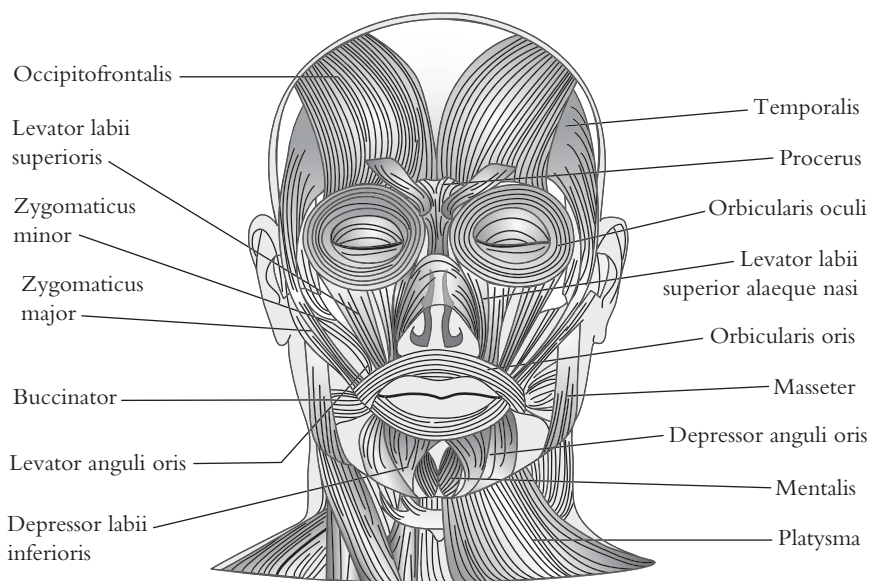
Figure 5.1: The main human facial muscles

5.1. Darwin on expression in humans and animals

Darwin's first edition of *The Expression of the Emotions in Man and Animals* appeared in 1872, a year after the publication of *Descent of Man* and a year

before the appearance of the sixth edition of *Origin of Species*. However, as he reports in his autobiography, he had been taking notes in the direction of such a work since the birth of his first child, William, in 1839 (Darwin 1969, p. 131–2). One main impetus of the project was to undermine the well known doctrine of Charles Bell that the human capacity for emotional expression was a gift conferred by God on man alone for the enhancement of social life. Instead Darwin would aim to show that expressive behavior, particularly as it occurs in the face, is widespread in the animal kingdom, and that much human expressive behavior is continuous with that of non-human animals. These two contentions would indirectly provide evidence that our expressive capacities are at least in part a result of natural selection rather than a divine gift. Darwin also aimed, by providing evidence for universals governing expressive behavior among human beings, to support the monogenecist (as opposed to polygenecist) thesis that modern humans are descended from a single ancestor (Darwin 1998, p. 355; Browne 1985).

Darwin holds facial expressions to manifest our emotional state, describing facial movements as revealing the state of mind of the organism (p. 356). He offers three main hypotheses that, in his view, jointly explain how this is possible: The Principle of Serviceable Habits, The Principle of Antithesis, and The Principle of the Direct Action of the Nervous System:

1. *The Principle of Serviceable Habits.* Certain behaviors are produced for the relief or gratification of internal states of the organism, and come to be habitually associated with them. As a result, those behaviors might be elicited by those internal states whether or not they are of use to the organism in every case in which they occur. Even when falling on a soft bed we often put out our arms to lessen the impact (Darwin 1998, p. 37). Rage brings about attack, and our ancestors bared their teeth to aid that attack. However, when now enraged we might bare our teeth even if no attack is in the offing, and this, for Darwin, explains why a sneer looks the way it does. The blinking characteristic of the startle reaction is another case, when it is clear that the organism is not in harm's way. Similarly, "I have caught myself, when thinking in the dark of a horrid spectacle, closing my eyes firmly" (p. 38). Because Darwin believed in the possibility of the inheritance of acquired characteristics, he also held that behaviors that

are habitually associated with internal states in one organism can become instinctively elicited by those states in its progeny.[1]

2. *The Principle of Antithesis*. As per Darwin's first principle, certain states of mind are associated by habit or instinct with certain behaviors whether or not those behaviors serve a purpose for the organism. What Darwin refers to as "opposite" states of mind consequently become associated with antithetical behaviors, again whether or not those behaviors are of service to the organism. Darwin gives as an example the contrast between a submissive dog, which will lower its head and tail, and an aggressive dog which will raise both its head and tail with hair bristling (p. 55). Pleasure and rage are in some sense opposed, and their characteristic facial expressions are noticeably distinct. Darwin apparently means by 'antithetical', behaviors that are perceptibly clearly distinguishable from others. Thus a lowered tail and an erect tail are antithetical in this sense because they are not likely to be confused with one another by the sorts of creatures needing to determine the affective state of the owner of the tail. Evidently, in spite of the suggestion of 'opposite', a single behavior can have more than one behavior antithetical to it.

3. *The Principle of the Direct Action of the Nervous System*. Darwin says little to elucidate this principle, but following authors such as Herbert Spencer he held that all animals with a central nervous system have a certain amount of "nerve force" that can be increased under the influence of strong emotion. In such cases distinctive behavioral consequences can follow. As an example, mentioned in Chapter 1, he gives the case of a boy whose hands shake so violently out of excitement that he is unable to reload his gun. So too, great pain tends to produce violent writhing of the body.[2]

---

[1] Darwin's examples might be called into question on the ground that the expressive behaviors he cites are of service to the organism in a probabilistic sense. That is, it might be held that putting out one's arms in reply to a sense of falling is *in general* a serviceable habit whether or not it is of service to the organism in a given case, and so forth for the other examples he considers. However, that claim is not plausible for the case of the sneer: It simply does not seem true that in general baring the teeth is of service to those feeling rage or resentment. That it may once have been the case that doing so was in general of service to the organism is no objection to Darwin's position here.

[2] Darwin appears to have derived this principle from Herbert Spencer, whom Darwin (1998, p. 74) cites as stating that it may be received as, "an unquestionable truth that, at any moment, the existing quantity of liberated nerve-force, which in an inscrutable way produces in us the state we call feeling, *must* expend itself in some direction—must generate an equivalent manifestation of force somewhere" (Spencer 1863, p. 109).

Darwin uses 'expression' in a wide manner so as to include behaviors that show emotions or other states of the organism whether or not they were designed to do so; the vignette of the boy and his shaking hands is a case in point. Many psychologists, ethologists, and evolutionary biologists follow him in this. We have given our reasons for adhering to a narrower usage, and to forestall confusion I shall sometimes speak, when discussing Darwin and those who work in his tradition, of manifestation or display of emotion rather than expression of emotion.

Darwin was fairly circumspect for most of his book about the communicative role of expression. Of the three Principles discussed above, only Principle 2, The Principle of Antithesis, suggests an explanation in communicative terms. Presumably the survival advantage that a species' conformity to this principle would confer is in facilitating conspecifics' determination of each other's emotional or other internal states, and Darwin suggests as much (pp. 63, 359). It is indeed difficult to see what other adaptive function adherence to this second Principle would serve.[3] Ekman points out that Principle 1, The Principle of Serviceable Habits, could conceivably be given as part of its rationale the communicative function of those habits, but Darwin does not offer this hypothesis. Only in the final chapter of his book does Darwin make some tentative general remarks about the communicative role of expression. There he writes,

There are no grounds, as far as I can discover, for believing that any muscle has been developed or even modified exclusively for the sake of expression … Nor can I discover grounds for believing that any inherited movement, which now serves as a means of expression, was at first voluntarily and consciously performed … On the contrary, every true or inherited movement of expression seems to have had some natural and independent origin. But when once acquired, such movements may be voluntarily and consciously employed as a means of communication. Even infants, if carefully attended to, find out at a very early age that their screaming brings relief, and they soon voluntarily practice it. We may frequently see a person

---

[3] It might be replied that the Principle of Antithesis does not need an explanation in adaptive terms. Rather, it might be held, it is inherently plausible in the way that, given a certain hydraulic picture of the nervous system, the Principle of the Direct Action of the Nervous System would be. This contention is dubious. No analogous model of the workings of the nervous system or any other part of an organism's physiology seems to make plausible the idea that distinct emotions must be manifested in discernibly distinct ways. The only plausible explanation that remains is the efficacy of this mechanism for communication.

voluntarily raising his eyebrows to express surprise, or smiling to express pretended satisfaction and acquiescence. (p. 351)

The first sentence quoted leaves open the possibility that some muscles have been developed in part for the sake of expression. However, the third sentence quoted above, beginning 'On the contrary', cannot plausibly be read in this manner. The 'some' occurring therein is most plausibly being used to mean the same as 'an exclusively' rather than 'a partly'. But on this construal, the movements in question are being said to have an origin that is entirely ''natural'' and independent of communicative purposes. It is only after that function is established (''But when once acquired'') that these movements can be used for communicative purposes. In the above remarks Darwin also shows his adherence to the possibility of expressive behavior as, at least in some cases, voluntarily performed. While the suggestion of his characterization of the smile is that it is not a sincere reflection of pleasure, that suggestion is not conveyed by his description of the raised eyebrows.[4]

Yet how can Darwin hold that expressive behavior did not evolve for the sake of communicating emotion, while at the same time holding that (a) at least some expressive behavior involves triggering innately given predispositions, and (b) some of this innately determined expressive behavior gives a competitive advantage to its possessor? I suggest the answer is that Darwin is implicitly thinking of innately determined expressive behavior as an *exaptation* rather than an adaptation. It has been suggested by ethologists that feathers on birds evolved because of their ability to keep the animal warm. However, a consequence of the presence of feathers was, once the feathers were sufficiently dense and covered enough area, to enable the bird to fly. Rather than being an adaptation, flight is thus an exaptation—a characteristic that gives its possessor a survival advantage but that did not evolve as a result of giving its possessor that advantage (Gould and Vrba 1982). Similarly, Darwin's view seems to be that while no expressive behavior evolved *because* it gave its possessors a survival advantage, its possession *does* in fact benefit that organism.

As an example consider Darwin's discussion of the human facial expression of grief. He holds that infants screaming from distress endanger their eyes because the vessels around them become engorged with blood. A

---

[4] See also Darwin's description of his own child's communicative development in Darwin 1877.

reaction that protects the eyes results in the contraction of the *corrugator* and inner portion of the *frontalis* muscles (which together produce furrows on the brow and forehead—see Figure 5.1), and that is precisely what produces a face disfigured with extreme distress. Darwin held that this protective reaction would have been learned in the first instance. It would then with repeated experience of this sort become a habitual response. Thereafter—thanks to Darwin's belief in the heredity of acquired characteristics—this facial configuration becomes an innate response in the organism's progeny. Here a facial configuration, even by the time it has become an innate reflex, is produced for the sake of protecting the eyes and not for communicating the grief that necessitates it. However, by virtue of the Principle of Serviceable Habits this facial configuration can be elicited by experiences of grief that are not so extreme that the well-being of the eyes mandates it (pp. 189–90). As a result, a facial configuration can become associated with grief of varying degrees,[5] and can thereby benefit organisms seeking relief from distress. Like feathers aiding in flight, it now serves a purpose for which it was not introduced. While Darwin's position, thus reconstructed, assumes the now-discredited possibility of inheritance of acquired characteristics, he could hold communicative behavior to be an exaptation without doing so.[6]

On this construal of Darwin, then, much expressive behavior is not communicative, and that expressive behavior which is did not evolve as a result of its communicative role. In this way Darwin puts severe limits on the explanation of expressive behavior in communicative terms. Burkhardt 1985 suggests that Darwin resisted explaining all expressive behavior in terms of communication because doing so would play into the hands of creationists in the tradition of Bell. If certain behavior has no communicative function then it cannot be seen as a gift from God for

---

[5] We assume here that while deception is of course possible, this facial configuration is not used deceptively so often that it ceases to be a reliable indicator of distress.

[6] Fridlund 1994, chapter 2, is one of the few authors who sees that Darwin does not hold that facial expressions evolved for the sake of communication; he cites many who do uncritically assume this on p. 15, fn. 2. However, Fridlund infers from this insight that facial expressions are on Darwin's view evolutionary vestiges, like male nipples in humans or webbed feet on land birds. This conclusion is undermined by passages such as those quoted above in which Darwin makes clear that facial expression does serve a purpose, as well as by the mere fact of Darwin's adherence to the Principle of Antithesis. The correct balance is achieved by viewing Darwin as holding facial expression to be an exaptation. This interpretation both accounts for actual text while giving Darwin a counter to Bell's creationism.

Table 5.1. Darwin's central tenets concerning facial expression

Are FE's signals?	FE's voluntary?	FE's inherently veracious?	FE's discernibly veracious?*	Are there pan-cultural FE's?	Distinction between ''felt'' and ''false'' displays?	Are some FE's non-social?**	Are FE's driven by an organism's strategic aims?***
Sometimes	Sometimes	No	Yes	Yes	Yes	Yes	n/a

* This is a telescopic way of asking whether it is possible for a normal observer to tell, of a given facial expression, whether it reflects a felt emotion or is instead dissimulating.
** This is the question whether facial expression occurs in the absence of an audience.
*** This question will arise for later theorists, but it is not one about which Darwin gives us enough information on which to base an answer.

the enhancement of social life; the remainder, which does have such a function, nevertheless did not evolve, at least in the first instance, for the sake of enhancing social life. The central tenets of Darwin's position are summed up in Table 5.1.

Many who have been influenced by Darwin's work on expression do not, for better or worse, share his anxiety to rebut creationism. Instead they have attempted in various ways to extend his ideas by emphasizing the communicative role of expression. We turn now to a prominent school of thought concerned with this task, the ''Neurocultural View'' of expression.

5.2. The Neurocultural View of facial expression

Proponents of the Neurocultural View hold, with Darwin, that expression may be seen as a manifestation of an emotional state, but, although they do not use this terminology, go beyond Darwin in suggesting as well that such manifestations have an evolutionary explanation in the specific sense of being signals. (Recall that this does not mean that they are intentionally produced.) Further, the Neurocultural View imputes a set of ''basic'' emotions—happiness, disgust, fear, surprise, anger, sadness being the most widely accepted—as each having its own pan-cultural facial signature. Accordingly, each of these facial signatures is held to be both signal and, because of its high correlation with a corresponding affective state, ''readout'' of that state. In our terms, when one of these characteristic

facial displays is in fact caused by its associated emotion, it both shows and signals that emotion.[7]

As an example illustrating the position of the Neurocultural View consider the smile. Some smiles are produced at will, such as those resulting from genuine though not overpowering pleasure over a current situation. Suppose for instance that you meet a friend whom you invite for dinner and who accepts your invitation. In hearing her accept your invitation you might be pleased in such a way as to have a genuine though not strong impulse to smile. The smile that may result would be polite, though not necessarily dissimulating. Other voluntary smiles include what Ekman and colleagues call the "miserable smile": Imagine a patient being told by her dentist that she needs to undergo a painful and costly procedure. One smile she might produce would convey a willingness to put up with the distress she is about to face, or an ironic delight over a depressing situation (Ekman and Friesen 1982).

In contrast to these two kinds of smile, a third has come to be known as the Duchenne smile, named in honor of the French physiologist G.-B. Duchenne de Boulogne, a contemporary of Darwin's. According to Ekman and colleagues, the Duchenne smile is a facial configuration produced in part by activation of both the muscle that orbits the eye (*obicularis oculi*) and the muscle that pulls up the lip corners (*zygomatic major*). The former muscle cannot be activated at will (Frank, Ekman, and Friesen 1993; Ekman, Friesen, and Davidson 1990), although in many cases the impulse for its activation stemming from the facial nerve can be inhibited. The Duchenne smile has also been shown to be associated with greater activation of the left frontal and left anterior regions of the brain as compared with other smiles. Ekman and colleagues also show that the congeries of muscular stimulation and facial configuration associated with the Duchenne smile is a highly reliable symptom of the occurrence of genuine enjoyment in the subject, as judged by self-report.

As might be suggested by its robust physiological basis, Ekman and colleagues have also made a powerful case for the view that the Duchenne smile is a human universal, part of our biological makeup rather than an exclusive product of culture. Similar views have been advanced for

---

[7] Other than the present work, I know of no defense of this approach that has considered whether the form of showing in question is showing-that, showing-α, or showing-how.

the other "basic" emotions such as anger, fear, surprise, sadness, and disgust. While not all of these emotions have been shown to have facial signatures that cannot be produced at will, or distinctive neurophysiological correlates, they are nevertheless presented as having a robust physiological basis that is independent of cultural variation (Ekman, Levenson, Friesen 1983). Insofar as cultural variation is to be found in the expression of these various emotions, this is due to "display rules" that have limited power in modifying and regulating their publication.

For just one example of the role of display rules, one group of Japanese and another group of American subjects were secretly videotaped while watching movies with both pleasing and disturbing themes, involving scenery and surgery respectively (Ekman and Friesen 1969). In the first experiment the subjects believed themselves to be unobserved. Quantifying the type and extent of facial movements in response to the films led to a high correlation (> .90) between facial expression in the two groups. However, when a scientist was brought into the room with the subject while the film was being viewed, facial expressions differed dramatically. The proffered explanation is that Japanese culture contains a display rule proscribing emotional display in front of an authority figure, whereas insofar as American culture contains such a rule, it is much less restrictive. It is assumed that these display rules are only activated when subjects take themselves to be being observed. When subjects believe themselves to be unobserved then culture is, as it were, disengaged and in general its display rules are inoperative.[8]

Thus characterized, it is a reasonable hypothesis that a facial expression such as the Duchenne smile is a signal in our technical sense: It may be elicited in response to environmental stimuli, and it is a reasonable hypothesis that in facilitating intraspecific communication about conspecifics' emotional state, it would give such a species a competitive advantage. For instance, whereas a threatening face might scare a competitor away from a piece of contested food without requiring battle, a smile or "playface" might appease a conspecific of higher rank, or encourage the approach of another for purposes of grooming or mating.

---

[8] These authors need not hold that culture is ever completely inoperative, and indeed it seems clear that people conform to display rules to some degree when in private. It is enough that there be a difference of degree that nevertheless leaves something in common among subjects of different cultures when unobserved.

At the same time, it is compatible with the Duchenne smile's status as a signal that it also be a genuine "readout" of an individual's affective state, and the correlation between occurrence of this facial configuration and self-reports of pleasure is high enough to justify viewing it thus. Accordingly Ekman and others have argued that at least to those observers with the requisite expertise for its detection, the Duchenne smile is a highly reliable symptom of the occurrence of an emotion. As a result the Duchenne smile is hypothesized to be both signal, in our technical sense, and "readout". Indeed as we shall see below, Ekman holds that a facial expression's status as a readout supports its ability to function as a signal.[9]

It should be clear that the hypothesis that a facial expression is a signal in the present technical sense of that word does not commit a theorist to the view that its production in a given instance is a case of speaker meaning. This is so even in terms of the notion of speaker meaning, bereft of communicative intentions, developed in Chapter 3. It is nevertheless a contention of the present study that human beings, and perhaps also higher primates, sometimes express themselves voluntarily by producing facial displays that are suitably analogous to involuntary manifestation of emotion. In this respect we agree with the quotation from Darwin given above in which he offers the case of a person voluntarily raising their eyebrows to express surprise. We thus do well to consider the merits of Ekman's (1997) apparently contrary view that facial expression is invariably involuntary.

In support of this position Ekman is careful to distinguish facial expressions of emotion from other facial movements whose primary role is to help guide conversation. These he calls conversational signals: "Here a facial movement is used much as the hands can be to illustrate speech as it is spoken. Facial movements, typically the eyebrows, accent, underline, or provide syntax for the speech as it is spoken." (1997, p. 330) I believe that Ekman means by 'syntax' here something akin to punctuation. The raising of the eyebrows at various points in a sentence may serve, for instance, as an unspoken exclamation mark. Although Ekman believes that many facial

---

[9] To the best of my knowledge proponents of the Neurocultural View of facial expression do not consider the possibility that it is an exaptation rather than an adaptation. I suspect that they could acknowledge this possibility without compromising other aspects of their theory.

conversational signals occur involuntarily, he grants that many also occur voluntarily.

Ekman also observes that from the fact that we can glean a great deal of information about a person from her facial expression, it does not follow that that expression was produced for the sake of conveying that information, and so was produced voluntarily. So too, Ekman points out, a person can experience an emotion without it manifesting itself on her face. Although various parts of the brain stimulate the facial nucleus, which in turn stimulates the facial muscles, most people are able to suppress the movements of those muscles to at least some extent. One reason that these movements are sometimes suppressed is that many cultures possess display rules, which govern the conditions under which various emotions may be displayed on the face and elsewhere. It is natural to suppose—and we have contended as much in Chapters 1 and 2—that at least for those people capable of suppressing facial displays of emotion, when an emotion's facial manifestation is not suppressed, its manifestation is one that we *allow*, albeit not one that we *will*. However, Ekman appears to challenge this line of thought, writing

The fact that expressions may often be managed by display rules, and that sometimes this management is voluntary, does not mean that the facial expressions of emotion that are being managed are also voluntary. If they were voluntary there would be no need to manage them. It is precisely because facial expressions of emotion are involuntary that we learn to manage these expressions, sometimes succeeding in totally inhibiting their visual appearance. The capability to inhibit an expression of emotion or modify it does not contradict my claim that the impulse for the expression is itself involuntary. (1997, p. 329)

To understand Ekman's position let us distinguish three things he might mean by 'involuntary':

First of all, a bit of behavior might be said to be involuntary in one "sense" just in case it results from the functioning of the Autonomic Nervous System.[10] That includes breathing when not governed by conscious control, pupil dilation, salivation, and sweating.

---

[10] For convenience, I speak here of "senses" of 'involuntary' without intending any commitment to the view that the word or any of its cognates is either ambiguous or polysemous.

A bit of behavior is involuntary in another sense just in case it is something that, at the time of its onset, we cannot help doing. Blushing is an example of this sort for most human beings, as are tears under situations of extreme distress: if the impulse to blush or weep is strong enough, there will be nothing—short of extreme measures like immolation—that we can do to prevent it.

Thirdly, an act or behavior might be thought involuntary when, although we can help doing the act in question, we cannot do it at will. I for one can't make myself sneeze, but I can prevent a sneeze so long as I notice its onset. Similarly, as we have seen, most human beings cannot produce at will the congeries of muscular configurations characteristic of the Duchenne smile, although this smile can often be suppressed. As a result, sneezing and the Duchenne smile are, in most cases, involuntary in the "can't will it" sense but not in the "can't help it" sense.

Being involuntary in the ANS sense does not imply either of the two other forms of involuntariness. Some things are involuntary in the former way while still being things we can prevent if we choose (such as breathing) and while still being things that we can do at will (again, breathing is an example). More generally, just learning that a bit of behavior is the result of ANS activity does not, on its own, show that it was involuntary in an everyday usage of that term. If we are to arrive at general conclusions relating facial expressions to voluntary or involuntary action, we need to see that it is only the second or third senses of 'involuntary' that correspond to our everyday notions of the involuntary.

We stipulated in Section 2.1.5 that the third sense of 'involuntary' given above is actually a form of voluntary behavior. There is little point in arguing about the propriety of a stipulation. What is more important is the question whether *any* of the above three kinds of involuntariness (or a proper subset of those three senses) justifies the thesis that only facial expressions in that sense of the term are sufficiently reliable indicators of emotions to show those emotions. It does seem plausible that facial displays that are involuntary in the "can't will it" sense are also expressions: since they can only have been produced by something other than the agent's conscious choice, it is a good bet that they also show the emotion that causes them. This does not imply Ekman's claim, however, for that claim, on the present interpretation, was that it is only such displays that are also

expressions. What about facial displays that we cannot prevent? Here again it seems plausible that it can show the emotion that causes it. On the other hand, could behavior that is not involuntary in either of the two senses just considered still show (and be designed to show) the emotion that caused it? [11]

Ekman seems to rest his claim for the involuntariness of facial expressions on the contention that if these expressions were voluntary, there would be no need to manage them by means of display rules. This is unconvincing, as an analogy with etiquette may show. Different cultures manage behavior at, say, the dinner table in different ways, and most oblige diners to manage their impulses in some way or other. This does not for a moment suggest that all etiquette-governed dinnertime behavior is involuntary, although of course some of it is. That conclusion would only ensue if all diners so internalized their culture's etiquette that for each norm enjoining doing A under condition C, they do A whenever C arises. Alas for the gracious host, that is not so. Likewise, it is plausible that the involuntariness of the impulse originating from the facial nucleus accounts for our need to inhibit some of our facial expressions. However, we have seen that in many cases, confronted with such an impulse we may decide whether or not to inhibit the facial display that it would issue in if left to run its course. That decision may take no more than a few milliseconds, but it does sometimes enable us to block the production of a facial expression if we so choose. Further, when we consciously choose not to do so, the result is a facial expression that we allow, even if it is not one that we will.

Ekman elsewhere argues for the hypothesis of the involuntariness of facial expression by contending that this hypothesis best accounts for the fact that facial expressions are reliable signals. He writes,

Not all signals are the same; emotional expressions are special, and we should not lose sight of that. They are special because they are involuntary, not intentional. Unlike 'A–OK' or 'good luck' hand gestures, emotional expressions occur without choice … The communicative value of a signal differs if it is intended or unintended. Emotional expressions have such an impact; we trust them precisely because they

---

[11] Even in such cases, of course, we can find moral responsibility. If I have a seizure while driving and kill a pedestrian in the process, there is no moment after the onset of the seizure and before it has run its course at which I can prevent it or the damage that it causes. However, I can still be held morally responsible for killing the pedestrian if, before getting behind the wheel, I could have been expected to know that I was prone to seizures.

are unintended…Our emotional expressions have been shaped and preserved by evolution because they are informative, but that does not mean that we deliberately make emotional expressions to signal information to others. (1998, p. 373)

These remarks come as part of a longer passage offered as what Ekman would like to say to Gregory Bateson, who at the time of Ekman's writing is deceased but with whom Ekman had debated these issues three decades earlier.

We may agree with Ekman's observation that just because emotional expressions have been shaped by evolution to be informative, it does not follow that we produce emotional expressions deliberately. Ekman's argument is nevertheless ineffective, and for two reasons. It is plausible that many emotional expressions are involuntary in either the ''can't help it'' or the ''can't do it at will'' senses, and thus indices in the parlance of earlier chapters. If so, then such facial expressions are highly reliable indicators of the presence of a genuine emotion. It may nevertheless be true that emotional expressions that are performed deliberately, are reliable indicators of emotion as well. Ekman gives us nothing to rule out this possibility. For all Ekman has shown us, the preponderance of facial configurations that purport to be manifestations of affect may indeed be such manifestations. Further, given the background of social norms mandating sincerity in facial displays, the likelihood that a given facial display is a display of a genuine emotion might be high enough to justify an inductive, though of course defeasible, inference from display to emotion without the further information that the display was involuntary. By the same token, while we do sometimes lose faith in the credibility of certain speakers, we often infer from what a person asserts to what she believes in spite of the possibility, in any given case, of dissimulation. This is particularly true in speech interactions in which agents know they are likely to meet each other again in the future and have an interest in maintaining their credibility. By comparison with the liabilities to credibility loss incurred by speech acts, our practices for keeping track of dissimulating users of facial expressions are relatively informal. Nevertheless we know perfectly well how to find someone phony in their use of their face. In spite of this, however, such things as smiles produced among members of a group with repeated encounters are often justifiably taken at ''face value'' even with the ambient possibility of dissimulation. Consequently, facial

displays can be reliable indicators of emotion without being involuntary in either the "can't help it" or "can't will it" senses of that term.[12,13]

None of our objections to Ekman's contention of the involuntariness of facial expression of emotion should be taken to carry over as objections to other insights provided by proponents of the Neurocultural View. This school has made a powerful case for the thesis that the approximately six basic emotions are displayed in ways that largely transcend cultural variation. They have also argued persuasively for the existence of robust physiological concomitants of facial displays. Further, proponents of the Neurocultural View have made a plausible case for the thesis that the capacity for emotional expression may be explicable in evolutionary terms, though in fairness to Darwin it should be conceded that their work thus far does not allow us to choose between the hypotheses that it is an adaptation as opposed to an exaptation.[14]

The Neurocultural View of facial expression also adds an empirical basis to the intuitive justification given in the previous two chapters for the view that some of our affective states are quite literally displayed on our face. For instance by isolating distinctive facial signatures for the basic emotions, the Neurocultural View helps us understand how an ordinary observer can discern, either in a Duchenne smile or a voluntarily produced display, the pleasure of the person whose face she sees. This is due to the fact that

[12] These observations build on and supplement those of Section 3.5 on the governance of speech acts by handicaps. The implication here is that facial expressions that can be produced at will may function as handicaps as well.

We also have enhanced memory for faces of those whom we believe to have been dishonest in the past; see Mealey, Daood, and Krage 1996.

[13] In fact, non-specialist observers are not particularly adept at distinguishing facial expressions that are produced deliberately from those that are not. This is so in spite of observable differences between these two kinds of facial expression. Continuing with our example of the smile, deliberate smiles also tend to be asymmetrical, and in particular tend to be more pronounced on the left side of the face (with right-handed subjects), whereas involuntary smiles are more often symmetrical (Hager and Ekman 1985). Duchenne smiles also involve activation of the musculature of the left side of the face prior to that of the right side, whereas this is not true of voluntary smiles. Nevertheless, Hess and Kleck 1994 show that observers asked to decide whether a given facial expression is spontaneous or deliberate do not perform better than chance. This is in accord with earlier studies such as Frank, Ekman, and Friesen 1993. In explaining their findings, Hess and Kleck show that observers use incorrect cues in distinguishing deliberate from spontaneous facial expressions.

[14] Ekman and colleagues have also provided evidence for what they call the Facial Feedback Hypothesis, namely that a voluntarily produced facial configuration often elicits the emotion (if there is one) typically associated with that configuration (Levenson, Ekman, and Friesen 1990). This Hypothesis is often considered part of the Neurocultural View, but we shall leave it aside because it is logically independent of the other tenets of this View and beyond the scope of the present study.

such a facial configuration is a characteristic component of the "affective program" of the so-called basic emotions, and we know from Chapter 4 that when X is a characteristic component of Y, relative to an organism's typical ecology, if such an organism perceives X, and Y is perceptible for that organism, it is also in a position to perceive Y. That an observer may be unable to rule out with certainty the possibility of dissimulation does not preclude her seeing pleasure when in fact the facial display is sincere. In this way experimental psychology fulfills the crucial task of determining whether our everyday experience is substantiated by experimental science. In the current instance the answer is in the affirmative.

Finally, the Neurocultural View, by situating facial expression in evolutionary perspective, helps to shed light upon the communicative status of this phenomenon. Human natural language is so strikingly different from the communicative systems of other animals that the explanation of its evolutionary development is a major challenge in contemporary theories of its origins. By contrast, human facial expression is continuous in many respects with that of nonhuman primates, and its evolutionary explanation does not seem inherently problematic: for facial expressions not governed by convention, we have a reasonable grasp of what would constitute evidence for or against a given evolutionary explanation, and the Neurocultural View's explanations seem plausible. It has nevertheless been argued that the Neurocultural View of facial expression gives a distorted view of the evolution and current communicative role of facial expression by failing to take into account the tenets of behavioral ecology. It is to this critique that we now turn.

5.3. The Behavioral Ecology View of facial expression

In recent years the Neurocultural View has come under attack from an approach to facial expression that not only emphasizes its strategic character as a signal, but also goes beyond this difference of emphasis to challenge its status as a readout. Heinroth nearly a century ago introduced the idea of *intention movements*: actions that are truncations of larger acts that sufficiently resemble those larger acts to be predictive of them. The baring of teeth sufficiently resembles the onset of a bite to be predictive of an impending

bite; as a result the baring of teeth can be a signal of an impending attack. Because the ability to give an indication of future action can be adaptive both for the sender and for the recipient of the message, a disposition to bare teeth in potentially agonistic situations can have survival value for both parties. Generalizing this idea, many authors including Andrew (1963a, 1963b) and Smith (1985) have suggested that expressive behavior be construed as a form of signaling in which the signal predicts with reasonable accuracy the organism's likely future course of action, and it may do so without being a truncation or icon of any such action. Thus a growl is not a truncation of an attack but is nevertheless a belligerent signal.

If the Neurocultural View may be caricatured with the word 'romantic', suggesting that emotional displays are emanations of what is within; this new line of thought may perhaps be given a cartoon representation with the word 'pragmatic', where what matters about emotional displays is not that they are a key to what is within, but rather to what the organism is liable to do in the future, and how effectively it will do it. Andrew sums up this distinction when he writes, "It is probably truer for a man to say, 'I would like to hit you' than for him to say 'I am angry.'" (1963b, p. 5). Andrew would presumably make a similar remark concerning the significance of a belligerent signal, namely that it is more accurately characterized as a predictor of future attack than as an emanation of rage. While Andrew's thought does capture a difference of emphasis between the romantic and pragmatic views of facial expression, we may nevertheless observe that it is not clear that Andrew's man needs to choose between the alternatives given. If an emotion such as anger has as part of its nature a tendency to attack, then being angry at someone just is, inter alia, to want to hit or otherwise harm them. This is precisely the picture offered by the view of emotions as "affect programs", and I shall return to it in a moment.

Fridlund refines this pragmatic approach to signaling, and in particular facial expression, with tools from behavioral ecology. Central to this school of ethological thought is the view that signaling evolves not only from the point of view of the signaler, but also from that of the recipient of the signal. We thus speak of *co-evolution* of communicative devices.[15] This will imply that selective pressure is exerted on displays so that they serve as signals that

---

[15] This perspective is elucidated further in Maynard Smith and Harper 2004.

are readily detectable by other organisms with which the signaler interacts; at the same time selective pressure is put upon the "receiver" organisms to develop and refine their means of detecting the "sender's" messages. Behavioral ecology also emphasizes the strategic character of signals. In this light, rather than seeing the bared-teeth display as a readout of anger, we see it as a signal that is sufficiently salient for observers to discern, and as sufficiently reliably coordinated with attack behavior, for an observer to draw the right conclusions from its production, for instance by retreating from disputed territory. By avoiding conflict both the signaler and recipient increase their chances of survival and reproduction, and thereby both act prudently. Fridlund offers the following scenario:

Millions of years ago, if you crossed my turf, I might bite your head off [at some risk to me, if you decided to retaliate]. If you had advance warning, you might escape death through retreat or protective defense, and we'd both survive. But you'd need cues to retreat or protect. I'd have to give them, and you'd have to notice them. Here's the scenario: because of a lucky gene, I adventitiously bared one tooth for $\frac{1}{2}$ second before I pounced. Your lucky gene made you look at my head. I bared my tooth, and you looked in the right place, not because I wanted to display my feelings, or because you wanted to see how I felt. We both acted out of pure dumb genetic luck. That we survived our skirmish increases the chances that our lucky genes will proliferate, and that my odd tooth-baring and your odd vigilance for it will both disseminate in our progeny. (1994, p. 76)

Fridlund thus emphasizes the complementarity of displays and vigilance for them, both intra- and interspecifically. For the interspecific case, it suffices to consider an example of Krebs and Dawkins: It behooves both bird and cat for the bird to signal, with a single wing flap, its ability to fly off before the cat can pounce on it. The production of this signal behooves the cat by preventing it from wasting resources stalking the bird, and it behooves the bird who will then not be interrupted from its feeding (1984, p. 388). Summing up the position, Fridlund writes,

For the contemporary ethologist or behavioral ecologist, facial displays are simply messages, which influence others' behavior because vigilance for and comprehension of signals co-evolved with the signals themselves.

Fridlund's central tenets include the following:

1. So-called expressive displays provide mutually beneficial signals of future action.

2. The dictates of economy and privacy would select against any involuntary displays of emotional information that would be detrimental to the displayer.

3. The costs and benefits of signaling would vary with the momentary social context and the animal's intentions within it.

4. Because facial displays are the results of a formalized co-evolution with vigilance for them, they are not readouts but tools that aid the negotiation of social encounters.

5. No distinction is made between "felt" and "false" displays issued by "authentic" and "social" selves—there is *only* a social self.

Fridlund proposes the Behavioral Ecology View of facial expression as a direct challenge to what we have been calling the Neurocultural View, which according to Fridlund harbors several fatal flaws. First, Fridlund contends that in focusing on involuntary facial expressions, the Neurocultural View ignores the costs of such expressions, particularly when betraying an organism's affective state can be dangerous for it. (Consider fear: It seems clear that an involuntary manifestation of fear is all but an invitation to attack, and an organism displaying fear is hardly likely to intimidate an opponent.) Secondly, according to Fridlund the Neurocultural View ignores the co-evolution of facial displays, focusing only on the sender of the message, rather than both sender and recipient. According to Fridlund, relatively little attention is paid among researchers in the framework of the Neurocultural View to, for instance, the discernibility of the emotions allegedly displayed on the face. Thirdly, the Neurocultural View treats voluntarily produced faces as dissimulating by contrast with, for instance, those of infants all of whose faces are presumably involuntary. In this respect it is simplistic, since as we have seen, an emotional expression is not, *simply* as a result of being voluntarily produced, a dissimulation. The main points of opposition between the Neurocultural and the Behavioral Ecology Views of facial expression are summed up in Table 5.2.

Recall the remark of Andrew 1963b that it would be more appropriate for a person to say, 'I would like to hit you' than to say, 'I'm angry at you.' Fridlund quotes this remark approvingly, and does so with good reason. For Fridlund and colleagues emotional expressions are not manifestations of what is within, but rather predictors of future action that are the result

Table 5.2. Darwin's, the Neurocultural, and Behavioral Ecology Views of facial expression compared

	FE's signals?	FE's voluntary?	FE's inherently veracious?	FE's discernibly veracious?	Are there pan-cultural FE's?	Distinction between "felt" and "false" displays?	Are some FE's non-social?	Are FE's driven by an organism's strategic aims?
Darwin	Sometimes	Sometimes	No	Yes	Yes	Yes	Yes	n/a
Neuro cultural View	Yes	Never	Yes	Only by experts	Yes	Yes	Yes	Yes**
Behavioral Ecology View	Yes	Sometimes	No	No	No	No	No*	Yes

* The Behavioral Ecology View does not deny that facial expressions are made in private, but contends that in these cases they are directed at a virtual or imaginary audience.
** The Neurocultural View sees facial expressions as being adaptive and thus as serving the interests of the organism on the phenotypic level; in this way they can serve the organism's strategic aims. However, such expressions, being involuntary, cannot be chosen as part of an organism's strategy. They can be so chosen according to the Behavioral Ecology View.

of co-evolution of both displayer and audience. We mentioned after our first reference to Andrew that it is not clear that one must choose between two ways of reporting one's anger, for anger just is, among other things, a tendency to violence. Hitting the object of one's anger is just the sort of thing an angry person feels like doing, and is what he will do in the absence of considerations of morality or prudence.

Similarly let us return to Fridlund's remark that, because facial displays are the results of a formalized co-evolution with vigilance for them, they are not readouts but rather tools that aid the negotiation of social encounters. This remark may be read in two ways, first as suggesting that the fact that facial displays are negotiating tools implies that they are not also readouts; the second as simply suggesting that the readout hypothesis is superfluous once the display-as-negotiating-tool view is accepted. The first reading leaves Fridlund with a position that is not compelling, for as we have seen it is far from clear that there is any conflict between facial displays being readouts and their being negotiating tools. Precisely because (at least) a "basic" emotion has as part of its nature characteristic tendencies to action, by manifesting such an affective state one thereby

manifests a disposition to act in a certain way. In so doing facial displays can aid negotiation by making clear an organism's likely response to attack, attempts at appeasement, or attempts at mating. Further, this picture is not essentially different for the more ''complex'' emotions; perhaps they too involve tendencies to action, but do so in a way more minutely related to other internal states such as cognitive states. Thus one who manifests hope manifests a tendency to be delighted at a certain outcome on the condition that she is aware of the outcome coming to pass; one who manifests contempt manifests a tendency to demean someone else under certain conditions.

The phenomenon is not limited to manifestation of emotions or moods. One who reveals her convictions thereby shows what she is willing to stand up for, and thus shows her tendencies toward future action, including but not limited to speech acts. More generally, in application to both cognitive and affective states, we see that the following may all be true at once:

1. An organism produces a facial display for the sake of manifesting its internal state.
2. That facial display shows the internal state the organism is purporting to manifest. If A shows B then B exists. Hence an organism may produce a facial display intentionally while showing a genuine internal state.
3. By showing a genuine internal state, the organism also shows its tendencies to act in ways characteristic of that state.

The first construal of Fridlund's claim does not seem tenable, so consider instead the second construal (that the readout hypothesis is superfluous once the display-as-negotiating-tool view is accepted). Thus interpreted, Fridlund's claim is also unacceptable once we see that the readout hypothesis buttresses the display-as-negotiating-tool view. It does this because it explains *how* a signal can be an indicator of future action. Some signals need not lead us to an affective or cognitive state of an organism in order to indicate its future course of behavior. Autonomic behavior is often produced in characteristic ways that lead to reliable predictions of future behavior. For instance rapid, repeated inhalation in the absence of physical exertion is an indicator of an impending sneeze. At another extreme, a contract between two parties can point to future behavior on the part of

both (that one will, for instance, deliver a set of goods while the other will respond with a payment) without also pointing to internal states such as intentions. The possibility of a contract pointing to future behavior depends upon the presence, mutually known by both contracting parties, of an enforcement mechanism. In the absence of an enforcement mechanism on the one hand, and a web of involuntary connections between a display and future action, on the other, a facial display of emotion signals future action by displaying our affective states, and thereby our tendencies to future action.

Let us next take up Fridlund's other dramatic claim, namely that there is only a social self. This seems to be motivated by the contention that no distinction is to be made between "felt" and "false" displays issued by "authentic" and "social" selves. In claiming that there is only a social self, Fridlund does not mean to deny that agents are sometimes solitary or even that they express themselves when they are. Rather he contends that when they do so their social life still suffuses their behavior. For instance, in response to the question whether all facial expressions are meant as negotiating tools, even when produced in private, Fridlund proposes that in this situation such expressions are intended for consumption by virtual or imaginary audiences. Fridlund thus vouchsafes an aspect of the behavioral ecology approach to facial expression by invoking a Goffmanian perspective on the self.

We may accept that at least organisms such as hominids, which have evolved in groups, are part of the social main. In this sense their selves are perhaps essentially social. However, this view of hominids does not for a moment undercut the thesis that such organisms are sometimes possessed of emotion, and sometimes not. Nor would it undercut the thesis that sometimes an organism purports to express an emotion it does not in fact feel. In such a case we have dissimulation. As a result we have yet to find a good argument from the thesis that at least certain organisms are essentially social, to the conclusion that the distinction between felt and false displays is specious.[16]

---

[16] Fridlund would, I suspect, ultimately reject the idea that there is such a thing as emotion, or better that the notion of emotion is rigorously defined. While the notion of an emotion is perhaps not well defined, we know significant things about emotions, and we are also sometimes justified in taking certain phenomena as explanatory primitives.

5.4. Strategic readouts: the face is a translucent strategist

Let us take stock. While showing that Ekman's arguments for the involuntariness of facial expression are—even if we accept their con‐clusion—compatible with facial expression's being voluntary in the, or at least an, everyday sense of this word, we have accepted other tenets of the Neurocultural View, namely that basic emotions are expressed in universally consistent ways with modulation by display rules, that they have important physiological concomitants, and that they can be revealed in the face. (Although Ekman does not do so, we glossed this notion of revelation in Chapter 4 to allow for some emotions being literally perceptible.) We have also agreed that it's a reasonable empirical hypothesis that certain facial displays such as the Duchenne smile are signals in our technical sense. At the same time we have seen the force of the ostensibly opposed Behavioral Ecology View of communication, a view providing a framework for seeing facial displays as signals that evolve as a product of co-evolution between signaler and audience, and which strategically aid both parties. While we may recognize the attractiveness of the behavioral ecology approach to facial expression, Fridlund's contention that facial expressions are not readouts is not justified, nor is his contention of the speciousness of the distinction between felt and false displays.

Fortunately, just as one can accept the other tenets of the Neurocultural View without agreeing that facial expressions are invariably involuntary (in one everyday sense), so too one can accept some central tenets of the behavioral ecology approach to facial expression while retaining a distinction between felt and false displays, and while retaining the view that some displays are genuine readouts of emotion. This fact provides us with the materials for fashioning a new position that preserves the best of the Behavioral Ecology and Neurocultural Views. Call this the Strategic Readout View (SRV) of facial expression. On the SRV, we agree that facial expressions can be genuine displays of emotion, indeed sometimes making those emotions perceptible, while accepting as well that facial expression, be it produced deliberately, or instead something that we allow to occur, is behavior governed by the tenets of behavioral ecology. Genetic tendencies to facial expression are results of co-evolution between the sender and

recipient of a signal. (Even if they are not strategically guided in the sense of being driven by a conscious plan, an organism's disposition to produce such an expression can be an adaptive trait, or instead an exaptation.) The SRV leaves a clear place for a distinction between felt and false displays, while agreeing that at least those organisms that evolve socially are themselves essentially, if not exclusively, social. Further, facial configurations that display emotion can signal future action precisely because what they display are, inter alia, tendencies to future behavior.

We mentioned at the end of section 5.2 that the Neurocultural View does the service of substantiating our common-sense belief that people sometimes wear their emotions on their face. The SRV incorporates this aspect of the Neurocultural View, but goes further in helping us to discern an empirical foundation for the approach to self-expression developed in earlier chapters. For the SRV helps us to see that it is not only true that we at times wear our feelings on our face, but also a plausible hypothesis that our tendency to do so is a result of co-evolution between senders and recipients of such signals. As a result, we may see facial displays as tools that aid the negotiation of social encounters, but we are now not tempted to infer that such displays cannot also be readouts; in fact it is their being readouts—signals that show what is within—that supports their utility in negotiating. At the same time, their being readouts is compatible with their being displayed voluntarily. The crucial points are summed up in Table 5.3.

The Strategic Readout View does not purport to intervene in all areas in which the emotions and Behavioral Ecology View might differ. For instance, Fridlund charges that the Neurocultural View fails to see that there would be selection pressure against involuntary displays that are detrimental to the organism. That may well be so, yet it does seem clear that animals do at times display emotion in a way that is detrimental to them. As a result, it is unclear whether a theory's postulation of involuntary, potentially damaging facial displays of emotion should be held against it, and the Strategic Readout View remains neutral on the question.

The SRV also does not presume to settle a difference between Darwin's position and that of the remaining two views on whether facial expressions are signals rather than exaptations in the technical senses of those words. Darwin, in holding that such expressions are exaptations, would be committed to denying that facial expressions are signals in our sense whereas the Neurocultural and Behavioral Ecology Views would contend that they

Table 5.3. Darwin's, the Neurocultural, Behavioral Ecology, and Strategic Read-out Views of facial expression compared

	FE's signals?	FE's voluntary?	FE's inherently veracious?	FE's discernibly veracious?	Are there pan-cultural FE's?	Distinction between "felt" and "false" displays?	Are some FE's non-social?	Are FE's driven by an organism's strategic aims?
Darwin	Some-times	Some-times	No	Yes	Yes	Yes	Yes	n/a
Neuro-cultural View	Yes	Never	Yes	Only by experts	Yes	Yes	Yes	Yes
Behavioral Ecology View	Yes	Some-times	No	No	No	No	No	Yes
Strategic Readout View	Yes	Some-times	No	Yes	Yes	Yes	Yes	Yes

are. The SRV does well to suggest that we will need to learn more from the fossil record of early hominid origins before this question can be settled.

Finally, it is now possible to explain how, as we incorporate the SRV into an overall picture of communication, we can in so doing also retain aspects of the extended senses model as promised in Chapter 1. Recall that the extended senses model suggested that in communication I serve as a prosthesis, or set of prostheses, for your senses; you do the same for me when you communicate with me. The attractiveness of this view needs to be set aside the fact that there is a difference in kind between my perceiving an object and your telling me about it; we also need to keep in view the fact that we care not just about those objects, but about one another's attitudes toward them. The SRV helps us meet the latter need. (Our account in Chapter 7 of how we show what our experiences are like will help us with the former.) In particular, as my face makes my emotion perceptible, I enable you to know how I am prone to behave vis-à-vis an object of common interest: My grimace shows the danger of your trying to get that object, or the need for your aid in defending against it. Such signals (or exaptive behaviors) are often strategically guided for the management of such social encounters, yet sensitive observers can "see through" them in

such a way as to discern the likely trajectories of one another's behavior. Facial expressions are translucent, then, in that we can peer through them into the likely course of one another's behavior; but we can also pay attention to those expressions themselves as we negotiate the social world. While the SRV may seem to apply most cleanly only to a small group of facial expressions, namely the expressions of the basic emotions, in the next chapter we will see that this view can apply even to cases in which we express ourselves in idiosyncratic ways, and even to those practices of self-expression that are subject to convention.

6

Convention and Idiosyncrasy

The varieties of self-expression discussed in the last five chapters have for the most part been non-conventional, and accessible to interpretation without specialized knowledge of the local culture or other idiosyncrasies of the expressing agent. Yet any satisfactory account of our topic must attend to the ways in which self-expression is subject to ritual, convention, and innovation. So too such an account must explain how expressive behavior can be at once immediate and visceral, while departing both from universal patterns of behavior be they convention- or biology-based. Our focus in this chapter is on how we show what's within in ways not typical of our species. In some cases such self-revelations will still enable what is shown to be perceived (6.2); that is so even when the expressive behavior is conventionalized (6.3). In yet other cases, discussed in 6.4, we show quantitative aspects of what's within, including our degrees of belief and the intensity with which we feel an emotion or mood. I explain how we do this by construing such forms of self-expression as types of measurement. These considerations prepare us for a fuller account of expressive devices in natural language, and I offer a case study of one such device in 6.5.

6.1. The limits of natural expression

As we observed in Chapters 4 and 5, many animals including human beings are biologically predisposed to facial and other forms of behavior that manifest their state of thought or feeling. The six basic emotions discussed in those chapters provide one such example in our own species: each of these emotions carries a characteristic facial signature and other behavioral markers that are relatively consistent across our species, and this fact seems due to our biological makeup rather than culture or choice. In addition, I have argued that facial expressions produced by these emotions are not

only correlated with them, but also characteristically show those emotions in the quite literal sense of making those emotions perceptible. (Such facial expressions may also show others how an emotion feels, and I will offer a hypothesis as to how they might do so in the next chapter.) As we also saw in those chapters, this predisposition to behavior manifesting what's within can be exploited by intelligent agents, who may produce or allow such behavior for the purpose either of manifesting their point of view, or giving others to believe that they are doing so, or both. In light of the Strategic Readout View of Chapter 5, we are now able to see as well that the actions of such agents, be they doings or allowings, may be at once strategically guided and capable of exhibiting genuine cognitive, affective, or other states of the self.

Once the exploitation of biologically grounded regularities has become possible, intelligent agents need not be bound by a requirement that their expressive behavior exactly conform to natural expressive patterns. Instead they might truncate, stylize, dramatize, or otherwise modify natural behaviors so long as in so doing it will be clear what aspect of their point of view they are trying to show. In fact this tendency to truncate, stylize, or dramatize behavior is already suggested by the Behavioral Ecology model, which, as we saw in the last chapter, emphasizes the interpretability of a display over its veracity (Fridlund 1997; Green 2003). Given that organisms are often in suboptimal communicating conditions—they may be in rapid motion, lighting may be bad, sound channels may produce distortion, attention may be divided, and so on—this is as we might expect. While we have found reason to object to the Behavioral Ecology model of facial expression on other grounds, the Strategic Readout View need have no quarrel with the idea that an organism would exaggerate or caricature a facial display if doing so aids in interpretability. A gain in interpretability in turn aids the agent in achieving her goal, namely acting on what's within by manifesting it.

Not only do natural forms of expression stand to benefit from modification for communicative efficacy, the very repertoire of those natural expressive forms can be limiting. For one, many states of thought and feeling seem to lack a natural form of expression. As artists well know, jealousy and disapproval are pervasive emotions lacking characteristic facial expressions or, apparently, other forms of stereotyped expression (Faigin, p. 14). So too with many other non-basic emotions: regret, pride, hope,

resentment, are just a few more cases. It is very hard, for instance, to show regret in one's face in the absence of a good deal of stage-setting. The same goes for trying to show regret in one's voice or other elements of demeanor in the absence of prop or convention (including semantic conventions).

So too for our more sophisticated cognitive states: Many such states also don't seem to admit of any natural expression. Even if it is possible to avail oneself of a natural expression of belief, it is quite unclear how one might express a conjecture or a supposition made for the sake of argument without recourse to some conventional device such as words ("I hereby conjecture that") or the subproof notation used in natural deduction systems. Further, we often have need to express not just the modality of a cognitive or affective state (that it is belief or anger), but also its content (belief *that* snow is about to fall, or anger *at* the unprovoked attack). Often those complex contents seem beyond the reach of natural forms of expression. We argued in Chapter 3 that one form of speaker meaning, factual speaker meaning, involves the manifestation of a state of affairs, and this suggests that some attitudinal contents can be shown, perhaps by ostension. However, many attitudes have as their contents facts that don't admit of ostension, and many others, being untrue, simply lack facts as their contents. For these reasons, if we are to express such contents, we need some ingenuity.

Certain stylizations of expression can also help achieve a purpose similar to the unmodified form but can do so more effectively. As we see from the Strategic Readout View, self-expressions are often produced with a strategic aim over and above that of manifesting the cognitive or affective state of their producer. A crucial such aim is the elicitation of a response from the addressee, for instance, succor for one in distress or intimidation in the object of rage. This is one reason why cries of distress may be replaced with mournful song, and why a face of anger may be accompanied by or even replaced by the brandishing of fists or other weaponry.

It is not difficult to see why brandishing (non-dental) weapons can be more effective in producing intimidation than a gnashing of teeth. Why should a mournful song be more effective than a cry of distress in producing succor? While a cry of distress may provoke distress in listeners by means of the well-known phenomenon of emotional contagion (Hatfield, Cacioppo, and Rapson 1994), such cries are generally not themselves engaging. Mournful song tends to be. While sad, a mournful song may also be pleasant to the ear simply by virtue of being a melody comprising notes

and rhythm. At the same time, if it uses such typical devices as minor keys, it will enable empathy in a way that simply perceiving someone's anguish will not do. We shall develop this line of thought further in Chapter 7. In the meantime, it will be well to consider further how self-expression is possible in the context both of idiosyncrasy and convention.

6.2. Expressive idiosyncrasy

In earlier chapters we saw two ways in which self-expression departs from natural forms of expression. First of all, for the non-overt case, we have seen that making as if to act on one's emotion is a way of expressing that emotion even when one has no intention of displaying it. In many such cases the mode of expression does not seem in any obvious way natural. Thus scratching out the eyes in the photo of a rival, or caressing a lock of a deceased loved one's hair, seem to express rage and affection respectively, yet neither seems naturally related to the emotion it expresses in the way that a scowl or a punch is related to anger. In contrast to these non-overt cases, overt displays involving pretense can also be expressive without being in any obvious way natural: You shape your hand into a gun, pretending to shoot at a person who has come late to a meeting. Here you show your anger by making as if to act on it, for by making as if to do so you show what you are inclined to do, and that inclination is itself a component of anger. If the mock shooting is done for the sake of making your anger overt, then you will have expressed as well as speaker-meant your anger whether or not anyone discerns it.

Showing-that and showing-α are both possible even when an agent is behaving idiosyncratically. While I can show my belief that P with an assertion having this very content, I can also do so by non-verbally behaving as if I believe that P; this I can do in myriad ways, many of which will in no interesting sense be typically associated with the belief in question. Likewise for many other introspectible states admitting of showing-that.

Consider now the case of showing-α. My expression of, say, emotion can be idiosyncratic even while it makes what is expressed perceptible. This is predicted by the account of showing what is within given in Chapter 4. For φ can be a characteristic component of A's ψ without being a characteristic component of ψ for every member of A's species. A's ψ

might show itself in distinctive ways due to A's personality or physiology. My beady look might show my surprise while your beady look might show your anger; the quaver in my voice is a characteristic feature of my trepidation while a quaver in yours might show your rising ire. Recall the passage from George Eliot considered in Chapter 2:

It was a moment of some agitation to both, though Philip had spent many hours in preparing for it; but like all persons who have passed through life with little expectation of sympathy, he seldom lost his self-control, and shrank with the most sensitive pride from any noticeable betrayal of emotion. A little extra paleness, a little tension of the nostril when he spoke, and the voice pitched in rather a higher key, that to strangers would seem expressive of cold indifference, were all the signs Philip usually gave of an inward drama that was not without its fierceness. (Eliot 2001, book VII, chapter VIII)

An idiosyncratic display of emotion such as Philip's enables perception of that emotion because it is a characteristic component of that emotion *for Philip*. Philip tends to show his agitation in his paleness, tensing of the nostril, and the rising pitch of his voice, and this could be true even if no other member of his species shows their agitation in this way. Even if that is the case, perception of this characteristic component of Philip's agitation enables perception of that agitation itself.

Perception is extensional: One can perceive these changes in Philip, and these changes can be characteristic components of Philip's agitation, with the result that one is also perceiving these characteristic components. What does not follow is that one knows that one is perceiving either these characteristic components or Philip's agitation. Knowledge, unlike perception, is non-extensional. (From the fact that I know that a is F, even if it is also true that a=b, we cannot infer that I know that b is F.)

What differentiates Philip's intimates from everyone else is that only the former know what they are perceiving when they see his paleness and tensed nostrils and hear his rising voice: Only his intimates know that what they are perceiving is his agitation. Others who see these same visible signs will also perceive his agitation without knowing that they are doing so. If, later on, they learn more about Philip they might reflect on the past and say, "What we were seeing back then was Philip's agitation, although we were too naïve to know it at the time."

Philip may not be among those intimates who know what they are seeing. That is, he may not be sensitive in the way that others are to the relations among his expressive behavior and his emotional life. Doing so would require careful observation of his behavior over a long period of time, and he might not have done this for reasons of lack of patience or interest or intelligence. Thus while Philip will generally know how he feels, it does not follow that he is more likely than others who know him well to perceive those feelings. Those who know him well may be better authorities than he on what aspects of his point of view he is manifesting.

What we have said about George Eliot's Philip applies not only to those who show what is within in idiosyncratic ways, but also to those who express themselves in ways that do not conform to universal or near-universal patterns. Philip's atypical manifestation of agitation does not appear to be intentional, or something that he allows for the sake of showing what is within; nor is it plausible that any other process, such as natural selection, has designed it to show what it does. His case is thus probably not one of self-expression. On the other hand, one can readily think of cases of a similar kind that are. To stick with the case of George Eliot, recall *Middlemarch*'s Rosamond Vincy, whose gentle twist of her neck shows her determination. In at least six passages in that novel, Eliot describes Rosie as twisting her neck slightly when she is feeling obstinate. Here is one:

This was a not infrequent procedure with Mr. Vincy—to be rash in jovial assent, and on becoming subsequently conscious that he had been rash, to employ others in making the offensive retractation. However, Mrs. Vincy, who never willingly opposed her husband, lost no time the next morning in letting Rosamond know what he had said. Rosamond, examining some muslin-work, listened in silence, and at the end gave a certain turn of her graceful neck, of which only long experience could teach you that it meant perfect obstinacy. (Eliot 2000, vol. I, ch. 36)

In this and other passages of a like kind, is it not clear that she twists her neck for the purpose of showing determination, or for the sake of making as if to act on that determination. It is compatible with the text that Rosie's neck-twist is no signal at all. However, one can imagine it becoming one. For instance, someone might point out to Rosie what

her twisting neck displays, and she might begin consciously to twist her neck for the purpose of displaying this "perfect obstinacy". If she does so, she may also be expressing her determination, and be doing so in such a way as to make that determination perceptible. In that case she will make her determination perceptible just as one might make one's anger perceptible by means of a scowl, or one's happiness perceptible by means of a (Duchenne) smile. What is special about Rosie's case is that her neck-twist is a characteristic component of her obstinacy. By contrast, my neck-twist is, to the best of my knowledge, not a characteristic component of my obstinacy.

6.3. The conventionalization of self-expression

We may now take Rosamond's quirk a step further by imagining not just that she uses it for communicative purposes, but that this behavior of hers becomes conventionalized. Imagine that Rosamond finds herself in an influential social position. People watch how she behaves and she is aware of this fact. Given this background, not only might Rosamond express determination by means of a neck-twist, others might follow her lead. As a result it may happen that the neck-twist propagates as a form of expressive behavior throughout her community in such a way as to achieve the status of a convention. It may, with time, even become automatic in the sense of that term used in Chapter 4.3, and thus become second nature, no longer requiring conscious intervention for its onset. Again, the ninth–tenth century Tibetan king Lang Darma was responsible for killing many Buddhists in that country. He was reputed to have a black tongue, and was much reviled. Sometime after his death people began to stick out their tongues at one another to show that theirs was not black, and, indirectly to show that they were out of sympathy with this tyrant. By now, the stuck-out-tongue gesture is used as an expression of humility.[1] It also seems to be a convention in certain parts of Tibet that one who sticks out one's tongue at someone expresses their humility. It is to the fuller account of the conventionalization of self-expression that we now turn.

---

[1] I am grateful to Frances Garrett of the University of Toronto for this information.

How can self-expression become conventionalized? I take a convention to exhibit, at the very least, the following three features:

1. A *regularity* in behavior
2. *Arbitrariness*: the regularity in behavior might have been otherwise
3. The regularity is supported by *normativity*: Given that all or most members of the relevant community conform to this pattern of behavior, for most members of that community, conforming to that convention is proper or appropriate.[2]

Examples of regularities are driving on the right side of the road, driving on the left side of the road, driving on the right side on odd-numbered days of the week and on the left side on even-numbered days. Other examples are putting knives on the right side of plates and forks on the left, using 'dog' to refer to dogs, sticking out your tongue to show humility, and pulling down one's lower eyelid with a finger to express skepticism. Further, each of these regularities might have been otherwise. Nothing about the end achieved by each regularity, if such there be, mandates that the regularity in question be the only one viable. One could either drive on the right or on the left, or even alternate. It doesn't much matter what the community does, so long as most of the community falls in, even if everyone in the community believes that their way of doing things is the only viable one.

The arbitrariness of a convention is compatible with its being 'hardwired' in a given species. Thus the fact that, say, a primate is biologically predisposed to use a certain sound pattern to alert others to a certain kind of predator does not rule out the possibility that this sound pattern has its meaning as a matter of convention. Even if it is hardwired, it might still have been otherwise. Likewise, a practice can be a convention even if everyone in the relevant community thinks it is not. Perhaps everyone in a certain society thinks their rules of etiquette are the only ones possible, or that their language is the only true language—all others being ''banging on a brazen pot'' in the words of Cratylus. That is compatible with their practice being a convention nonetheless.

---

[2] A convention also seems to require that parties to it not only do, but also intend to coordinate their activities as per the regularity in question—otherwise we would have a coincidence of behavior without a convention. I do not, however, know of a satisfactory formulation of this mutuality condition, so I leave it out of the present account.

Finally, about normativity. Moral imperatives, etiquette, practical rationality, theoretical rationality, are all sources of norms. These are sources of such norms as "keep promises", "send thank-you notes for gifts", "maximize utility", and "believe only what is true", respectively. Further, for present purposes I need not take a stand on such vexed issues as whether moral normativity is a species of practical normativity; for all the present discussion requires, each source of normativity might be independent of all the others. That said, what I mean in saying that conventions yield norms is that a convention makes conforming to the regularity in behavior that it comprises the, or a, "right" way to behave. Utensil-placing conventions imply what is the, or a, right thing to do in setting the table. Linguistic conventions imply what is the, or a, right way to refer to dogs (for users of the relevant language). Expressive conventions, as we shall see, imply what is the, or a, right way to show skepticism (for members of a certain culture), humility, contempt, and so on.

In light of this brief gloss of the notion of convention, we may now ask what it is for self-expression to be conventionalized. Here is a minimal condition: *an expressive convention must yield a proper way in which one can show one's introspectible state.* As we saw in Section 6.1, in the absence of conventions, our expressive repertoire is limited by the fact that some of our introspectible states either cannot, or cannot readily, be perceived or otherwise shown. This is true both of the modality of such states, and their content. So too the limitations of communication channels might make "natural" forms of expression relatively inexpedient. For this reason it behooves a group to institute a practice by means of which its members can show an aspect of themselves that cannot, or cannot efficiently, be perceived or otherwise shown.

There are limits to what a convention can do. If a given state of thought or feeling cannot be perceived, then there could also be no convention bringing it about that this state is perceived. Given our discussion of Chapter 3, we know that if such a state is to be perceptible, that is either because some characteristic component of it can be perceived or because it is possible to perceive that state either in or through some other medium. Thus for instance it is not clear that I can show my gratitude or jealousy in such a way as to make it perceptible. We have expressions in English that are conventional expressions of gratitude, such as 'Thank you,' but if gratitude cannot be perceived then no convention has the power to

change that fact. Further, even if gratitude can be perceived, it does not follow that use of a certain form of words, or one's disposition to use such words, is a characteristic component thereof. What, then, does an expressive convention do, given that all self-expression involves showing what is within?

Self-expression is a way of showing, and expressive conventions would enable us to show states of ourselves. Demonstration is not specific to self-expression. I might demonstrate, and thereby show, the rising numbers of women in poverty as a result of the recent tax cuts. Or my calculations on the chalkboard might demonstrate the presence of a binary star at a certain point near the Eagle Nebula. These are demonstrations not in the logician's strict understanding of that notion (of a sequence of syntactic objects each of which is related to the previous according to deductive rules), but rather in its everyday understanding. According to such a quotidian understanding, I might demonstrate my love for someone by sacrificing something of great value (an ear, a career) to be with her. When I do that I have shown my love as well. I have thereby given part of myself in the hopes of receiving something in return. Expressive conventions go a step further by enabling their users to show the presence within them of certain states with a mere gesture, speech act, or other conventional device rather than with a material sacrifice. Like other forms of demonstration, they take us from barter to currency, but as with other uses of currency, one who issues an expressive "coin" still gives up something, namely a portion of commitment-freedom. One who renounces a measure of commitment-freedom is, in effect, agreeing to a handicap in the sense of that term elucidated in Chapter 3.

My presentation of a check for $100 purports to demonstrate my ability to pay the payee that amount when the promissory note is cashed. Whether or not it does demonstrate that, my presentation of the check does show my *commitment* to paying when the note is cashed. In making that commitment I give something of myself, that is I undertake a handicap; in particular I give up a commitment-freedom on this issue with the result that I am now obliged to pay under certain circumstances. As a result I close off certain possible avenues of action, such as spending the entire balance of my funds on something else with no danger of another's retaliation. In issuing the promissory note, then, I forfeit something. In so doing, I also show my commitment to paying under certain conditions.

In a similar manner, in issuing an apology I forfeit something, also a measure of commitment-freedom. I am now committed to regretting or feeling remorse for the action for which I am apologizing. Given this forfeiture, together with the being in force of a convention governing apologies, my credibility as an issuer of apologies, and my actually feeling regret, I show that regret. In forfeiting a bit of commitment-freedom with the proximate aim of showing my regret, I hope to gain your appeasement in return. You might show this appeasement to me with such words as, "Fuhget about it!"

The institution of expressive conventions, then, creates a kind of currency that can be spent to procure social goods, while at the same time admitting the possibility of counterfeit. I wish to make amends for some infraction against another, and I invoke the convention of using 'I'm sorry' to apologize. If accepted (with words such as "Fuhget about it!") this coin buys me the appeasement of the party I've wronged. In an insincere apology, I offer a counterfeit coin which might achieve my aims as well as the genuine article. Such insincerity does, however, put me in danger of being found out, and if that happens my ability to use this expressive "currency" in the future may be impaired. I would have a reputation for passing wooden nickels, expressively speaking, and the result would be that it becomes difficult for me in the future to use this convention (of saying 'I'm sorry') to show remorse or regret even when I do feel it.

Expressive conventions, then, take the following form:

To show one's Ψ, do Δ,

where Ψ is a state that is also a possible object of introspection. Thus, as charmingly discussed in de Jorio 1832, Neapolitans of his day follow a hypothetical imperative to the effect that

To show your skepticism, pull your eyelids apart,

In much of the English speaking world (at least), playgrounds are governed by the hypothetical imperative to the effect that

To show your contempt, stick out your tongue,

whereas in Tibet we have the norm,

To show your humility, stick out your tongue.

These are what we might call inclusive, rather than exclusive norms. They do not tell us what are the only permissible ways of showing what is within, but rather tell us what are some permissible ways of doing so. As with many other conventions, the institution of expressive conventions does not require conscious deliberation by anyone. A given expressive convention might have been brought about by a conscious choice, but need not have been. Instead such conventions as the above might have arisen as unintended consequences of large patterns of behavior. Whether they have done so in any given case is a question for a historically oriented linguistic anthropology.

When we express ourselves in ways that enable perception of what is shown, we do not thereby commit ourselves to being in any particular state. I show my determination in my face in such a way as to enable you to perceive it; there is then no further question of sincerity. By contrast, the intentional employment of convention-invoking showing creates a commitment on the part of the agent so acting. One who says, 'Oi', 'Ouch', or 'Yuk' is committed to being annoyed, in pain, or disgusted, respectively. In the absence of such emotional or experiential states being felt by the agent, she cannot be said to be sincere, and if found out she will be subject to a loss of credibility. This suggests the following thesis: What I forfeit when I invoke an expressive convention is a measure of freedom from commitment. When I invoke that convention I thereby give up some of that freedom in that I am now committed to being in a certain affective or other internal state. Whereas perception-enabling showings, and convention-free demonstrations, are the barter of emotional life, convention-involving demonstrations are more sophisticated transactions whose currency is commitment. One who pulls down her lower eyelid, shrugs her shoulders, or sticks out her tongue, if acting with the intention of invoking a relevant convention, is committed to being skeptical, indifferent, contemptuous, or humble (depending upon the cultural milieu), respectively. In so doing she has laid down her cash (rather than her cow).

How can the above conventions, which, as I have argued, enable one to show what is within, be used in aid of self-expression? This latter phenomenon is a matter of producing a signal that shows one's introspectible state, and the above conventions are social artifacts whose job it is to enable us to show such states. This is what makes them expressive conventions.

Expressive norms may govern words as well as gestures and sound patterns. 'Ow' is used as an expression of pain; 'Yuk' as an expression of disgust; 'Oi' an expression of annoyance. I have seen groups of people among whom a calmly uttered 'Grrr' is an expression of anger. Notice that all these expressions are bereft of semantic content. None of these words seems to refer to anything, to express a property, to behave as a connective, or to possess any other semantic feature typical of contentful expressions. They nevertheless have a conventional meaning. Accordingly, such words have a conventional meaning but not because they possess a semantic content. Instead they possess a pragmatic prowess; call it an *expressive meaning*. That meaning is mastered by one who knows the expressive norm governing these words, and part of this mastery consists in the knowledge that such words can be used to show one's state of thought or feeling. (This notion of expressive meaning is elucidated further in 6.5 below.)

In addition to words that have no content but do have an expressive meaning, natural languages such as English seem to contain words and expressions that have both a semantic content and an expressive meaning. 'Idiot', 'knucklehead', 'pinhead', 'moron' all have a semantics: they are nouns that are true of an object just in case it is not intelligent. However, to assert of a person that he is an idiot, knucklehead, or the like is among other things to derogate that object, to express an attitude of condescension or contempt for that object's intelligence. The reason is that these words are governed by a convention to the effect that they may be used to show one's contempt for that object's intelligence. (In this respect they are more articulate than a word like 'Yuk'; they place restrictions on what can be the object of one's contempt, whereas 'Yuk' places no restriction on how it can be used to direct one's disgust toward something.) That has not always been the case for 'idiot' and 'moron', which were once expressively neutral scientific terms. However, these words have by now acquired a sense that we may characterize in expressive terms. Further, recalling our remark in Chapter 2 that it is possible to express an attitude by means of a ''saying in one's heart'', one who thinks to herself, ''That moron,'' in reference to an item of her attention is expressing her contempt for that person without publicizing that contempt. She is still showing that contempt, albeit only to herself.

Expressive linguistic devices are not found only among words or expressions. For instance grammatical moods may have as part of their meaning an

expressive norm. Similarly, logicians have contemplated the introduction of a sign indicating the logical status of the sentence to which it is prefixed. Frege, for instance, proposed an assertion sign for a logically correct language. A sign such as this would evidently be a conventional indicator of commitment to the proposition it prefixes. Similarly, grammatical structures such as parenthetical expressions seem to have an established expressive dimension, namely indicating the nature of the speaker's commitment to some other part of the proposition in which the parenthetical expression occurs.

In some cases of self-expression, the state expressed is made perceptible. In Chapter 7 we shall see that in certain other cases what is shown is not perceptible but rather is something with which another appropriately sensitive being can empathize. In such cases our self-expression enables others to empathize with us, and thereby shows others how some aspect of what is within feels. Typically, conventionalized self-expression enables neither perception of what is expressed nor empathy with what is expressed. There are, however, cases in which conventionalized forms of expression are evocative, or, as an aesthetician might say, express*ive*. This way of describing an expressive act suggests that it comes closer than others to being in some sense iconic of the attitude it is conventionally used to express. Imagine a culture as similar to ours as is consistent with the further stipulation that instead of an extended middle finger, people express extreme contempt by overtly touching their earlobe. That practice would seem less visceral than our own. The reason seems to be that one can ''see in'' the extended middle finger a violation of a person's sexual privacy, while no analogous imaginative reconstruction seems forthcoming in respect to the touching of the earlobe. I shuddered with quotes in the last sentence because we do not literally see any sexual violation in an extended middle finger. Rather, in thus seeing the extended digit we readily imagine it to be a penis or other instrument of sexual intrusion, and we may further readily imagine it, as it is thrust toward the object of someone's contempt, being at liberty to violate that person. The dramatically extended middle finger, then, depicts its audience as subject to sexual violation and therefore as an object of contempt. At least in Western cultures, by contrast, nothing of this order can readily be ''seen'' in the touching of an earlobe.

As another example, Koch 1983 remarks that the hand-wave suggests a striving toward the audience of that wave; this is naturally read as the view

that the hand-wave mimes an action of striving toward the audience of that wave. This in turn is naturally construed as an action that is readily imagined to be the waver herself striving toward her audience. Another convention could use another gesture to greet or bid farewell, but this one seems apt precisely because it is iconic of the action that we are inclined to perform when we greet or bid farewell.

Just as we can "see" in some conventional expressions acts or relations that appear to show an attitude, in other conventional devices we might "hear", that is imagine hearing, or otherwise sense a person's attitude in a tone of voice or, for that matter, in a searing guitar riff. Nothing in principle rules out imaginative perception through other sense modalities besides sight and audition. Because, however, of this role of the imagination required in the observer of a conventional expression, it would not be plausible to hold that the observer can literally perceive the emotion thereby expressed. We do not, in general, see or otherwise perceive what we make as if to perceive. Even when expressive, then, conventional forms of expression allow us to show what we think or feel without making it perceptible.

6.4. Measuring what's within

Conventionalized self-expression shows what is within in a way that depends on the existence of a system of conventions relating internal states to publicly accessible tokens. This form of showing is found not only in the area of self-expression, but also with such devices as speedometers and bar codes. In such cases what is shown is not an internal state of an organism, but that does not upset our analogy. What is crucial, rather, is that, for instance, the numeral on the speedometer shows the vehicle's speed; in so doing the numeral need not make that vehicle, or its speed, or its moving at a certain speed, perceptible; and further that the numeral on the speedometer shows the speed in spite of its not being infallibly correlated with the vehicle's speed. When it fails in this way, the speedometer is malfunctioning and does not show the speed of the vehicle.

We remarked in Chapters 1 and 2 that not only can a person express anger, pain, or fear, she can also express states that have contents. In some cases I might express anger without expressing the content of that anger. (I might just say, 'Grrr' without showing what the anger is about.) In

other cases, we express not only a state but also its content. That might be achieved by my expressing a content in a certain way. For instance, my son might not only refer to my daughter, but speak her name angrily. In so doing he is expressing not only anger, but also expressing anger *at Sofia*. Again an utterance of an indicative sentence might not only express a content (usually the semantic content of those words—exceptions may be malapropisms and slips of the tongue), but might also express an attitude toward that content by expressing that content with conviction. A satisfactory account of self-expression needs to make sense of this duality in what can be expressed. In the rest of this chapter I shall do so by proposing the following position: natural language not only has a characteristic role of representing states of affairs; it also has a characteristic role of showing the internal states of its users by serving as a measuring system for those states. Attitude ascription generally, I shall suggest, is analogous to measurement, and a speech act (at least those that express an internal state) is analogous to self-measurement in which I show, if I am both reliable and sincere, my state of mind—in particular both its modality and content. Given that I'm pretty reliable about asserting only what I believe, when I sincerely state that the sun has just gone down, I not only mean that the sun has just gone down, I also show my belief that the sun has just gone down, and I do so not in such a way as to make that belief perceptible, but in a way that is convention-involving. The two crucial conventions are, first, those that give my words meaning, and, secondly, the grammatical and intonational cues that give my utterance assertoric rather than some other force.

The analogy between attitude ascription and measurement stemmed, historically, from the desire to address a worry about the metaphysical significance of intentionality. Must attitude ascription be seen as putting us in contact with such abstracta as propositions? Without attempting to resuscitate adverbial or sentential accounts of such ascriptions, some recent authors have contended that an agent's believing that P no more requires her to stand in a relation to P than does an object's weighing five pounds require it to stand in a relation to the number 5. So long as the chosen measurement scale keeps track of the relevant empirical relations among objects (such as balancing on an equal arm balance) it does not matter which particular numbers are used as indices of those relations. Similarly, so long as a system of attitude ascription keeps track of such phenomena as an agent's preferences among outcomes or dispositions to assent, it does not

matter which propositions are used to index them. Because the choice of a number to index a magnitude (proposition to index thoughts or behavior) is within certain limits arbitrary, the urgency of the idea that in having a magnitude an object bears a relation to a number (in thinking or behaving thus and so the agent bears a relation to a proposition) is lessened. In spite of their relational surface grammar, sentences reporting attitudes need not, according to this line of thought, be given relational truth conditions. By challenging the question's presupposition, this conclusion may exonerate us from answering the question how standing in a relation to a proposition can be causally efficacious.[3] It nevertheless bids fair to make sense of how one's assertion of P shows one's belief that P (so long as we agree, as we should, that a reliable thermometer might show the outside temperature to be 45 degrees Fahrenheit).

Some definitions will ease formulation of the analogy between measurement and attitude ascription. Say that *U is a relational system* iff U is a finite sequence $< A, R_1, \ldots, R_n >$, where A is a (non-empty) set of elements (called the *domain* of U) and $R_1, \ldots, R_n$ are relations on A. Where $U = < A, R_1, \ldots, R_n >$ and $U' = < B, S_1, \ldots, S_n >$ are relational systems, *U' represents U* iff U' is a homomorphic image of U, that is, iff there is a function f from A into B such that for all $x_1, \ldots, x_k$ in A, $R_i(x_1, \ldots, x_k)$ implies $S_i(fx_1, \ldots, fx_k)$. U is a *numerical relational system* iff the domain of U is a set of real numbers, and U is a *formal relational system* iff its domain consists of abstract objects such as numbers, points, sets, and so on. Once a relational system has been given a formal representation, the question arises to what degree that representation is unique. Where a formal relational system $U' = < B, S_1, \ldots, S_n >$ represents relational system $U = < A, R_1, \ldots, R_n >$ by means of the mapping f, a transformation T from B onto itself is admissible with respect to f just in case $T[f(x)]$ also bears witness to the representation of U by U'. The set of all admissible transformations determines the scale type or the degree of uniqueness of the system of measurement chosen. (Among the better known scale types definable on numerical relational systems are *nominal scales*, on which all one–one transformations are admissible, *ordinal scales*, on which all

[3] The analogy between attitude ascription and measurement has been defended by Churchland 1979, 1989; Leeds 1979; Field 1981; Stalnaker 1984; Davidson 1974, 1989; Swoyer 1987; and Matthews 1994. Churchland 1979 advocates an adverbial analysis of attitude ascriptions, but as Davidson 1989 remarks, this untenable position is logically independent of the measurement analogy.

monotone transformations are admissible, *interval scales*, on which all positive linear transformations are admissible, *ratio scales*, which are uniquely determined except for an arbitrary unit of measurement, and *absolute scales*, on which no transformations are admissible (an example is counting).)[4]

It is natural to hold that where T is an admissible transformation of f, then if $f(\mathrm{x}) \neq \mathrm{T}[(f(\mathrm{x})]$ the particular entity onto which f maps x is to some extent an artifact of the system of representation chosen. This is perhaps why assigning the number 5 to an object as its weight need not be understood as placing it in a relation—for instance, the weight-in-pounds relation—to that number. If it were to be so understood then by virtue of weighing-in-pounds what it does the object would stand in a relation to a great many other numbers as well. Such a view of measurement would also raise the question how an object's having the weight-in-pounds that it does, construed as its standing in a relation to an abstract entity, could be causally efficacious.[5] Rather, following Suppes and Zinnes 1963 we may say that what is, or at least purports to be objectively correct in any system of formal representation is what is common to all admissible transforms thereof. For instance, instead of holding that a non-absolute measurement scale relates objects to particular numbers, we may hold that any such scale ascribes features to objects by indexing these features with abstract entities that preserve the significant aspects of those objects and their interrelations.[6]

Those who contend that more than one system of intentional ascription can capture all the significant features of an agent's thought and speech acts may see an analogy here. Suppose, for instance, that only one of two empirically adequate systems of attitude and speech act ascription imputes to an agent the belief that P, the other ascribing no belief with that content. One of these systems of ascription may nevertheless be related to the other in a manner analogous to that in which one measurement system is related to another that results from it by means of a scale transformation. If so,

---

[4] For further development of these concepts see Suppes and Zinnes 1963.

[5] The point is well made by Crane 1990, pp. 225–6.

[6] Stalnaker writes, ''What is it about such physical properties as having a certain height or weight that makes it correct to represent them as relations between the thing to which the property is ascribed and having a number? The reason we can understand such properties—physical quantities—in this way is that they belong to families of properties which have a structure in common with the real numbers. Because the family of properties which are *weights* of physical objects has this structure, we can (given a unit, fixed by a standard object) use a number to pick a particular one of the properties out of the family'' (1984, p. 9).

then since the two systems of intentional ascription differ over whether the agent in question believes that P, the ascription of that belief must, one might suggest, be an artifact resulting from the choice of one system of ascription over the other. It is not *false* to say of the agent that she believes that P, any more than it is false that her body temperature is 98.6 degrees Fahrenheit. Rather, both ascriptions contain elements that are, strictly speaking, innocuous but able to tempt us into unwarranted ontological conclusions.

In helping us to discern the line between the factual and the artifactual, the study of admissible scale transformations and possible interpretive schemes helps to guard us against such temptation. In the temperature case we may clearly exhibit that line by characterizing the set of admissible transformations for any system of temperature measurement, and an attempt to delineate the distinction between the factual and artifactual in attitude ascription may also be of value. For if we are mindful that all that purports to be objectively correct in any system of attitude ascription is what is common among all empirically adequate rivals, we will not be tempted to infer that the agent stands in a relation to the proposition P just from the fact that one empirically adequate system of attitude ascription imputes to her an attitude with that content. At the very least that inference awaits the finding that all empirically adequate systems of ascription do so. Davidson has put the point this way:

The same facts may be represented by quite different assignments of numbers. In the interpretation of speech, introducing such supposed entities as propositions to be meanings of sentences or objects of belief may mislead us into thinking the evidence justifies, or should justify, a kind of uniqueness that it does not …. Propositions being much vaguer than numbers, it is not clear to what extent they are overdesigned for their job. (Davidson 1974, p. 147; see also his 1980, p. 6, and 1989, pp. 9–11.)

The analogy between attitude ascription and numerical representation (hereafter *the measurement analogy*) is unlikely to admit of quick validation or refutation. For one, unlike in the case of, say, the measurement of mass, we have no concise and uncontroversial account of the relations on the empirical domain that belief–desire ascription is intended to preserve. This lacuna need not vitiate the measurement analogy—it was, after all, only in 1901 that Hölder explicitly characterized the empirical properties and

relations that the measurement of mass is to capture—but it does hamper its assessment. How, then, might the analogy be developed?

In an early discussion of the analogy between attitude ascription and measurement, Field (1981) develops a suggestion of David Lewis's that the domain of the empirical relational system that attitude ascription may be seen as representing comprises internal states having sentence-like properties. One virtue of this approach is that it suggests an account of what properties and relations in the empirical domain are to be preserved in the process of representation. On this approach, if for instance internal state S, in virtue of its syntactic properties, syntactically implies internal state S', in virtue of *its* syntactic properties, then any formal system representing this relation by means of a mapping f must be one in which $f(S)$ logically implies $f(S')$. If the domain of the empirical relational system is construed as a set of sentence-like entities, then, as Field suggests, we will have some guidance in determining the relevant properties in the empirical relational system requiring to be preserved in representation.

Field takes these considerations to support a view of the internal states to be represented by a formal relational system as having sentence-like properties, but other proponents of the measurement analogy have demurred. Stalnaker (1984, chapter 1), for instance, argues that the relations in the domain of the empirical relational system that are to be preserved in any formal representation are causal and functional characteristics of those states. One of those states, which would be represented in the formal relational system as the desire that P, disposes its possessor to act in such a way that she brings it about that P in a world in which her beliefs, whatever they might happen to be, are true. Another of these states, which would be represented in the formal relational system as the belief that P, disposes its possessor to act in such a way that she satisfies her desires, whatever they might happen to be, in a world in which P (as well as her other beliefs) are true.[7] It does not seem, however, that such functional characterizations of these states require imputing to them linguistic structure, and Stalnaker argues at length (1984, chapter 2) that such an imputation is not compulsory.

---

[7] Stalnaker recognizes that while this account might serve to determine that a given internal state has the attitudinal type of belief, it is not adequate to determine that state's content. What is needed is not just a functional account of the connection of internal states with behavior, but also an account of their production by states of the environment (Stalnaker 1984, pp. 18–19). As will emerge presently, Matthews's development of the measurement analogy respects these constraints.

Matthews 1994 develops a similar line, contending that the domain of the empirical relational system that attitude ascription may be seen as representing comprises those internal states of organisms that are causes of their optimizing behavior. On Matthews's account, the principle guiding our mapping of those states onto the representational space used in attitude ascription is that the properties preserved in the mapping, "are those causally efficacious properties of internal states in virtue of which possessors of these states usually succeed in behaving optimally in their environment, i.e., in a way that satisfies their needs" (pp. 139–40). He infers from this that to a first approximation an internal state that is shaped by the environment is represented by the system of belief–desire psychology as belief, whereas one that shapes the environment is represented by that system as desire.[8] We ascribe contents to those states as a way of indexing the distal environmental situations that shape and are shaped by them; characterizing an internal state of an organism as a belief that P, then, tells us that it is one that is typically brought about by the state of affairs P, while characterizing an internal state of an organism as a desire that Q tells us that it is a state that typically brings it about that Q. The representational system available from belief–desire psychology does not itself incorporate any postulate linking attitudes and action, but it may be used in conjunction with such logically independent postulates as that agents strive to fulfill their desires given that the world is as they take it to be, to allow users of that scheme to rationalize behavior.

We need not here take a stand on the dispute between Field, on the one hand, and Stalnaker and Matthews on the other, concerning the amount of sentence-like structure to be imputed to items in the domain of the empirical relational system that is represented in attitude ascription. So too we may remain neutral on the question of the exact structure of the formal relational system that will represent attitudes. All proponents of the measurement analogy agree that the formal relational system must contain

---

[8] For further development of this approach see Dretske 1988, particularly chapters 3 to 6. We note also that on the present formulation it is not the case that the items in the empirical domain are assumed in the process of formal representation to be beliefs or desires. It might turn out that any adequate representation of those items by means of a formal relational system will map them onto entities having belief-like or desire-like properties, but that conclusion would be the upshot of a general account of successful representation and not an initial desideratum for any adequate formal representation. Consequently, in supposing that the items in the empirical relational system are beliefs and desires (from which he infers that their belief-like properties, such as having the contents that they do essentially, must be preserved by any system of formal representation) Crane (1995, p. 194) mischaracterizes a desideratum for a successful articulation of the measurement analogy.

propositions, but they will differ over how much, if any, internal structure must be imputed to propositions. Just to fix ideas, however, let us observe that Matthews (1994, pp. 135–6), following Richard's (1990) synthesis of Russellian and sententialist conceptions of attitudinal content, contends that the space onto which items in the empirical domain of belief–desire psychology are mapped comprises ordered pairs $< a_i, < s_j, r_k >>$, where a_i is an attitude type, s_j a sentence-type and r_k a Russellian proposition (itself an ordered n-tuple consisting of $n-1$ individuals and a property (or relation)) capable of being expressed, relative to a context of utterance, by s_j. This approach allows the items used to index attitude states to individuate them at least as finely as do sentences, while still representing those states as contentful. Observing that not all aspects of a representation space can be read back into the empirical domain that it represents, Matthews (pp. 138–9) leaves it open whether a function mapping internal states of organisms into a space comprising propositions justifies the conclusion that all inferential relations among propositions have an image in relations of internal states to one another. Consequently, such failures in agents as departures from logical consistency in beliefs and from deductive closure in beliefs do not provide an immediate objection to the empirical adequacy of belief–desire psychology as here construed.

Matthews does not state, much less prove, a representation theorem for belief–desire psychology, perhaps because it is insufficiently clear which relations among internal states this form of psychology is intended to capture.[9] On the question of admissible transformations, Matthews argues

---

[9] Following the suggestion above concerning the empirical properties and relations to be preserved in representation by a formal relational system, a simplistic statement of such a theorem might take the following form. The empirical relational system (ERS) would have the form $<U$, *tendstobringabout, tendstobebroughtaboutby>* where U is a set comprising both internal states of the organism and states of affairs in the organism's environment, the latter conceived as the standing in a relation R of finitely many objects, $o_1, \ldots, o_n$. We shall refer to the state of affairs of relation R holding of $o_1, \ldots, o_n$, with the expression '$R(o_1, \ldots, o_n)$'. 'Tendstobringabout' and 'tendstobebroughtaboutby' name relations that are borne between states both of the organism and of the environment. (One is not the converse of the other: whereas the belief that P tends to be brought about by the state of affairs of P's obtaining, it is not the case that that state of affairs tends to be brought about by any belief.) We will accordingly have such empirical relations to represent by means of the formal relational system as *tendstobringabout*$(a, R(o_1, \ldots, o_n))$, and *tendstobebroughtaboutby*$(a, R(o_1, \ldots, o_n))$, where the first relation is characteristic of desire and the latter of belief. The formal relational system (FRS) will have a domain comprising the set of real numbers R, as well as a set of Russellian Annotated Matrices, or RAMs such that for every state of affairs $R(o_1, \ldots, o_n)$ in the ERS there is a RAM in the FRS having the $n + 1$-tuple $< o_1, \ldots, o_n, R>$ as its second element and as its first element a sentence S capable of expressing $< o_1, \ldots, o_n, R>$ in some context of utterance. The FRS will also contain the binary relations BEL and DES, defined

that whether two representations are of the same propositional attitude depends on whether, in ascribing an attitude, our explanatory interests require us to be faithful to the ascribee's words or point of view; he concludes that an account of the degree of uniqueness of the system of representation employed by belief–desire psychology must invoke pragmatic considerations. However, this view is not required for adherence to the measurement analogy, even after the formal relational system's domain has been characterized as comprising propositions as conceived by Richard. In an attempt to characterize the space of admissible transformations one could instead advert to the indeterminacy confronting, say, Davidson's radical interpreter. From such a perspective an account of the degree of uniqueness of attitude ascription need not invoke pragmatic considerations. Rather, on such an approach only a criterion of empirical adequacy need be invoked in determining whether one system of attitude ascription is to be construed as an admissible transform of another.

Consistent with other authors who have taken up the measurement analogy, Matthews conjectures that attitude type will always be preserved by admissible transformations: even if one system of attitude ascription imputes to Smith the belief that P while an admissible transform of that system ascribes to him the belief that Q, any admissible transform of either

on pairs whose first element is a real number and whose second element is a RAM. Among the relational facts that might hold in the FRS will thus be $BEL(k, < S, < o_1, \ldots, o_n, R>>)$ and $DES(k, < S, < o_1, \ldots, o_n, R>>)$, where k is a real number, S a sentence, and $< o_1, \ldots, o_n, R>$ a proposition that S expresses in some context of utterance.

With the aim of representing the ERS with the FRS, we begin with a mapping f that takes internal states into numbers, and states of affairs to RAMs. For the internal states, we use Axiom of Choice to ensure that the set of internal states has a well ordering, and we number all the states thus ordered starting with 1 and going upward. To map propositions to RAMs, observe that there will in general be many RAMs having the same second element, and a function from a proposition in the ERS to the RAMs in the FRS must choose among these. Again, order all the RAMs having the same second element by means of Choice, and define f as follows:

Where $R(o_1, \ldots, o_n)$ is a state of affairs in the domain of the ERS, define $f(R(o_1, \ldots, o_n))$ as the first RAM according to the foregoing ordering whose second element is $< o_1, \ldots, o_n, R>$.

For a proof that FRS represents ERS according to our definition of f, it needs to be shown that as f has been defined,

if *tendstobringabout*$(a, R(o_1, \ldots, o_n))$, then $DES(f(a), f(R(o_1, \ldots, o_n)))$ and if *tendstobebroughtaboutby* $(a, R(o_1, \ldots, o_n))$, then $BEL(f(a), f(R(o_1, \ldots, o_n)))$.

We need not attempt such a proof here.

system will construe the internal state in question as a belief. Matthews is right to enter this as a conjecture rather than assume it as self-evident. After all, it is not self-evident that the only causally efficacious properties of internal states in virtue of which possessors of these states usually succeed in behaving optimally in their environment can be the properties of being beliefs or desires. Among other candidate properties are those in virtue of which agents may be seen as maximizers of utility.[10]

We have pursued the analogy between attitude ascription and measurement far enough to see its value. When I ascribe an attitude to someone else, with such words as 'John believes that P,' I relate him to a formal relational system that, if accurate, shows what he believes (desires, etc.), without making his state of mind perceptible. It shows, in particular, the content of that state by mapping it into a space of propositions, and its modality (that it is a belief rather than a desire, conjecture, or supposition for the sake of argument) by mapping it into a system of forms of commitment. (Belief, conjecture, and supposition for the sake of argument are all forms of commitment to a proposition, but they carry with them strikingly different internal norms as elucidated in Chapter 4.) Moreover, when I perform a speech act for which there is a sincerity condition (such as asserting, promising, thanking, or conjecturing, and unlike appointing), I implicitly relate myself to such a relational system. How? Suppose I assert that P. Then I invoke a set of conventions according to which I can be represented as bearing the belief-relation to P. This is not crucially different from the way that a triage nurse in an Emergency Room might ask me to rate my pain on a scale of 1 to 10. I might just touch the number 6, and thereby express (and thus show) my pain as well as its degree without making either perceptible. Attitude ascriptions that are not avowals, such as second- or third-person ascriptions, or past-tense ascriptions to myself, are like the nurse saying, of me, 'He is at 6.' Attitude ascriptions that are avowals, by contrast, are like my pressing the '6' button from among a scale of 1 to 10.

---

[10] Accordingly, it cannot be assumed without argument that different ways of representing mass are analogous to different systems of ascribing beliefs and desires. Unless it can be shown that beliefs and desires are the only internal states that can account for their possessors' optimizing within their environment, a less tendentious view is that different ways of representing mass are analogous to different ways of ascribing *attitudes*, which may or may not require to be described in terms of the concepts of belief and desire. I argue for this in greater detail in Green 1999c.

The visible disgust on my face shows my state of mind and, indirectly, the noxiousness of what I have just tasted. Similarly, my sincere and correct assertion of P shows my state of mind if I am honest, and, if I am both reliable and honest, the obtaining of the state of affairs that P. In contrast to the disgust case, my assertion does not make my state of mind perceptible. Rather it shows the content of my state of mind by virtue of the semantic conventions governing my words, for those conventions allow us to map the meanings of those words onto a space of propositions. It shows the modality of my state of mind (that it is a belief rather than a conjecture or a wish) by virtue of pragmatic conventions governing my utterance, for those conventions allow us to map that utterance onto a distinct space containing various modes of commitment.

On the approach we have offered, then, self-expressions are still readouts, and they are still characteristically produced for strategic purposes. For this reason the account here sits naturally with the Strategic Readout View of Chapter 4. Readouts that invoke convention, in contrast to those we considered in Chapter 4, typically do not enable perception of what they display. Because, however, such readouts often involve mapping my internal state onto a relational system with revealing internal structural properties, such readouts can be more effective in inquiry, planning, and exchanging information than self-expression of the more ''brute'' variety.

6.5. Some verbal devices of self-expression[11]

In this section I offer a case study of an expressive device in English, the parenthetical construction. This device is well suited for illustrating our account of how self-expression can be conventionalized, and indeed so thoroughly embedded in everyday communicative practice that we use it unreflectively to show the attitude we take toward other things we say. To see why, observe that even after disambiguation, assignment of context-sensitive contents, and other pragmatic determinations of what is said, speakers often fail to convey the force with which their indicative

---

[11] This technical section may be skipped by readers more concerned with questions of aesthetics taken up in Chapter 7, which does not presuppose any of the current section's conclusions.

sentences are uttered. Dialogues such as the following are for this reason not uncommon:

(1) α: You'll be more punctual in the future.
     β: Is that a prediction or an order?

Speakers sometimes preface their remarks with clarificatory material to forestall questions such as β's. α might, for instance, have prefaced his words with 'This is a prediction.'[12] However, because this prefatory remark is itself an indicative sentence and thus might be put forth with other forces than assertion, prefacing one indicative sentence with another free-standing indicative sentence cannot be relied upon to clarify the illocutionary force of the former. It would thus behoove speakers to employ a force-elucidating device whose own illocutionary status is not itself up for question.

   Some writers concerned with this problem have suggested that certain constructions in natural language *show* rather than *say* what the force is of a remark to which they are appended.[13] On a suitable character-ization of the distinction between saying and showing, this claim may permit a view of certain expressions as reliably indicating the force of the remarks to which they are appended. However, it is by no means clear that saying and showing are mutually exclusive activities. For instance, parenthetical expressions often serve to indicate the nature of

----

   [12] A propos this topic Strawson writes:

The speaker … not only has the general authority on the subject of his intention that any agent has; he also has a motive, inseparable from the nature of his act, for making that intention clear. For he will not have secured understanding of the illocutionary force of his utterance … unless his complex intention is grasped. Now … for the enterprise to be possible … he must find means of making the intention clear. If there exists any conventional linguistic means of doing so, the speaker has both a right to use, and a motive for using, those means. One such means … would be to attach, or subjoin, to the substance of the message what looks like a force-elucidating *comment* on it, which may or may not have the form of a self-ascription. (1964, pp. 451–2)

As examples of force-elucidating comments, Strawson has in mind cases such as 'This is a warning' and 'This is only a conjecture' prefacing indicative sentences.

   [13] Urmson, for instance, discussing parenthetical expressions such as 'as I conjecture' and 'as I fear', writes:

parenthetical verbs are one of the sets of devices that we use in order to deal with these matters [of clarifying the illocutionary force of one's remarks], though not the only set. By them we prime the hearer to see the emotional significance, the logical relevance, and the reliability of our statements. This we do not by telling him how we are moved or how he should be moved by them … but by the use of warning, priming, or orientating signals; we show rather than state. (1952, p. 197)

a speaker's attitude or commitment toward a proposition she has elsewhere expressed, and they do so in such a way that their own illocutionary status does not need to be made clear. Consider, for instance,

(2)   If snow is white, as I believe, then grass is green.

One who asserts (2) makes clear among other things that she believes that snow is white, and as we shall see in more detail presently her parenthetical remark can only be taken as expressing her commitment rather than, say, conjecturing it or querying it. On the other hand, by means of that parenthetical remark the speaker also *says* that she believes snow to be white, and the speaker would have said something different had she instead uttered

(3)   If snow is white, as John used to think, then grass is green.

As a result, just to invoke a distinction between saying and showing does not suffice to account for how some ways of showing express commitment reliably whereas others do not. We require an account of how some ways of showing can reliably indicate the force of the comments to which they are appended in spite of having semantic content of their own, which content is essentially capable of being put forth with any of a variety of illocutionary forces.

6.5.1. *Strong illocutionary force indicators*

To clear the ground for such an account, let 'Δ__' be, syntactically, a function from sentences into sentences, chosen from a set C of connectives such that each element of C is in the domain of a function IF, whose range comprises illocutionary forces. This allows us to speak of 'the force associated with connective 'Δ__''. Then let us say that 'Δ__' is a *strong illocutionary force indicating device* (hereafter *strong ifid*) iff for all sentential complements S, the utterance of '$\Delta(S)$', outside quotation marks, is the putting forth of S with the illocutionary force associated with 'Δ__'. Let us also adopt the policy that utterance of a sentence that contains '$\Delta(S)$' as a proper part is also an utterance of '$\Delta(S)$'. This definition is thus meant to include embedded as well as free-standing occurrences of the uttered expression (quotation marks excepted). It is not clear that there are, or even could be, any strong ifids in an actual or possible language. Davidson 1979, for instance, has argued that no expression could be a strong ifid, on

the ground that any expression purporting to be a strong ifid could be used by actors on stage to heighten the realism of their performances; in such uses, however, no speech act is performed even though the actor may be making as if to perform a speech act.[14]

6.5.2. Illocutionary force and illocutionary commitment

Davidson's challenge to the possibility of a strong ifid is cogent but does not show the notion of force indicator to be theoretically useless. For another notion of force indicator not susceptible to Davidson's criticisms is available. With a view to developing such a notion let us observe that being committed to a semantic content A under mode Δ is not to be identified with performing a speech act having force Δ and content A. One who asserts A, for instance, is committed to all that follows logically from A even though it is doubtful that she has *asserted* all those propositions. As discussed further below, we do have good grounds for taking assertion to distribute over conjunction, so that one who asserts A&B asserts A and asserts B. However, because asserting A seems to require in the very least that the speaker intend to be committed to A,[15] it does not seem that one can be said to assert unrecognized logical consequences of what one asserts. For this reason it does not appear plausible that assertion is closed under logical consequence. On the other hand, lest it be impossible to criticize someone who asserts propositions that jointly imply, say, a contradiction, we must preserve the idea that one is *committed* to all that follows from what one asserts. The notion of illocutionary validity is designed to track such commitments. This notion allows us to capture the idea that one who asserts A is committed to all that follows from A, but without our thereby implying that the speaker has also asserted any of the propositions deducible from A. Similar remarks apply to commitments to contents that arise from speech acts other than assertion.

Not only is commitment to a proposition distinct from assertion of that proposition, commitment to a proposition that one has asserted is one among various sorts of propositional commitment. As we saw in Chapter 3, assertion of A differs from, for instance, a conjecture that A in

---

[14] As argued in Green 1997, neither Hare 1989 nor Dummett 1993 succeeds in countering Davidson's challenge.

[15] Our account of illocutionary speaker meaning of Chapter 3 bears this out.

that one who asserts that A is liable to be challenged to provide adequate grounds for believing what he asserts. Often we make such a challenge with the words, "How do you know?" In contrast, it is inappropriate to respond to one who merely puts forth A as a conjecture with the question, "How do you know?" While decisive grounds for the falsity of A do require one who conjectures A to retract that conjecture, her inability to respond to a challenge to establish A does not oblige her to relinquish her conjecture. Just as the nature of commitment undertaken in assertion of A differs from the nature of commitment undertaken in conjecture that A, so too it appears that one who asserts A is committed to A's logical consequences in a manner different from the way in which one who conjectures A is committed to its logical consequences. For reasons such as this, in propounding the notion of illocutionary validity it is insufficient merely to describe speakers as committed to propositions and other semantic contents; accuracy requires also tracking the mode of that commitment by adverting to the force of the speech acts that engendered it. The concept of illocutionary validity is designed to track the mode as well as the content of speakers' commitments.

6.5.3. *Illocutionary validity and weak illocutionary force indicators*

Let S be an arbitrary speaker, $< \Delta_l A_l, \ldots, \Delta_n A_n, \Delta B >$ a sequence of force–content pairs; then:

> $< \Delta_l A_l, \ldots, \Delta_n A_n, \Delta B >$ is *illocutionarily valid* iff if speaker S is committed to each A_i under mode Δ_i, then S is committed to B under mode Δ.[16]

Let V_S be a verb taking a sentence nominalization as complement, and where S is a sentence let S* be the nominalization thereof, namely 'that S'. Let '$[[\delta]]^c$' denote the semantic value, relative to context of utterance

---

[16] This definition is close to the definition of weak illocutionary commitment given in Searle and Vanderveken 1985, chapter 4. An argument is illocutionarily sound just in case it is both illocutionarily valid and all its premises are such that their conditions of satisfaction are met. Developing a related notion of Kearns's (Kearns 1997), say that 'ΔB' is illocutionarily necessary iff every sequence of force–content pairs having 'ΔB' as its terminal element is illocutionarily valid. Although it is beyond the scope of the present study to provide one, a fuller account of illocutionary validity would employ further distinctions. For example, we could distinguish among the different possible objects of commitment, since there is nothing to rule out being committed to a question or to an imperative. See Harrah (1980, 1994) for discussion of assertoric, erotetic, and projective commitment.

c, of δ. Let '__, as NP V$_S$,' be, syntactically, a function from sentences to sentences. We observe that 'as' in '__, as NP V$_S$,' is a prosentence, anaphorically dependent, typically, on an S that occurs in the argument place of '__, as NP VP,'.[17] Consequently, where this prosentence is anaphorically dependent on S$_i$, $[[as]_i$ NP V$_S]]^c = [[$NP V$_S$ S$_i^*]]^c$. Let '⊢' denote assertoric commitment. Then we say that '__, as NP V$_S$,' is a *weak indicator of assertoric commitment* just in case for all illocutionary forces f and sentences S, the inference

$< f \ [[\ldots S_i \ldots, as_i$NP V$_S, \ldots]]^c, \vdash [[$NP V$_SS_i^*]]^c \ >$

is illocutionary valid. According to this inference, a speaker committed under any mode to the semantic value, relative to context of utterance c, of a sentence in which 'S$_i$ ..., as$_i$ NP V$_S$,' occurs, is also showing her assertoric commitment to $[[[as]_i$ NP V$_S]]^c$, that is, to $[[$NP V$_S$ S$_i^*]]^c$. (Her doing so is compatible with her putting forth S with other forces as well. It is also compatible with her intending any of various perlocutionary effects.) For instance, '__, as John contends,' would be a weak indicator of assertoric commitment just in case anyone committed under any mode to the semantic value, relative to context of utterance c, of a sentence in which 'S, as John contends,' occurs, thereby shows her assertoric committed to [[John contends S∗]]. (Hereafter where there is no danger of confusion I shall omit the qualification 'relative to context of utterance c' as understood.)

As-parentheticals are sometimes used to convey an even further layer of illocutionary commitment. For an utterance of

If, as I believe, P, then Q,

commits the speaker not only to the conditional, if P, then Q, and to the claim that she believes that P, but also to P itself. In uttering this sentence in a speech act she shows her assertoric commitment to P. (This claim does not follow merely from the fact that she shows her assertoric commitment to 'I believe that P': She might become committed to that simply on the basis of evidence gathered by a clinical psychologist that she believes, say, that walking under ladders is bad luck. She might in this case become committed to 'I believe that walking under ladders is bad luck' without thereby putting forth, 'Walking under ladders is bad luck'

---

[17] The word 'as' is thus a proform of the same category as 'true' and 'so'. Grover, Camp, and Belnap 1975 formulates a view of these words as proforms taking indicative sentences as antecedents; Grover 1992 further develops the approach.

assertorically.) This is due to a conspiracy of two facts, the first of which is that all as–parentheticals, when used in a speech act, illocutionarily entail the corresponding ''slifting'' construction, that is, for all forces f

$$< f \ [[\dots S_i \dots, as_i NP \ V_S, \dots]]^c, \vdash [[S_i, NP \ V_S]]^c >$$

Secondly, some slifting constructions illocutionarily entail their complement with the force signaled by the parenthetical element, that is, some slifting constructions participate in the following illocutionary validity:

$$< \ \vdash [[S_i, NP \ V_S]]^c, f_\Delta \ [[S_i]] >.$$

For instance, an assertion of 'P, I believe' is inter alia an assertion of P. We can see this clearly by noting that one who asserts 'P, I believe, but it is not the case that P' has literally contradicted herself. Similarly, an assertion of 'P, I guess' is inter alia putting forth P as a guess but not as an assertion. These cases are in contrast to 'P, John claims' which just attributes a claim to John with no commitment to the truth of P as well. All as–parentheticals illocutionarily entail the corresponding slifting construction, but only some slifting constructions illocutionarily entail their complement. Where V_S participates in an illocutionary validity of the above form, where '＿, as NP V_S,' is a weak indicator of assertoric commitment, and where f_Δ is the mode of commitment characteristic of verb V_S, V_S thus generates the following *derived* illocutionary validity: for all illocutionary forces f and sentences S, the inference

$$< f \ [[\dots S_i \dots, as_i NP \ V_S, \dots]], f_\Delta \ [[S_i]] >$$

is illocutionarily valid. I shall call any illocutionarily valid inference having this form a *weak-ifid elimination inference*, and will call any expression of the form '＿, as NP V_S,' and participating in a weak–ifid elimination inference a *weak ifid*. Were there any weak ifids, then although their presence in an utterance would not purport to answer the question *whether* the utterer is performing a speech act, it would provide at least a partial answer to the question *what* commitments the utterer is undertaking and under what mode.[18] Finally, let us say that a weak ifid is *robust* if and only if it has embedded occurrences.

---

[18] We are now able to see, therefore, that the notion of a weak ifid is not circular. Also, Davidson's skepticism about the possibility of a strong ifid does not carry over to the possibility of a weak ifid. For further discussion see Green 1997.

6.5.4. *Parenthetical attitudinatives as robust, weak ifids*

We are now in a position to argue that certain parentheticals containing attitudinatives or speech act verbs are robust, weak ifids. (A related thesis concerning non-restrictive relative clauses also is defensible, but the elaboration and defense of this further thesis would take us too far afield.) As an example of the first of these consider

(4)   If snow is white, as I suppose, then grass is green.

One who asserts (4) is: (i) claiming that if snow is white then grass is green, and (ii) showing her assertoric commitment to snow's being white. (Assume that 'suppose' here means 'believe' rather than 'suppose for the sake of argument'.) The latter act is performed even though the parenthetical clause is embedded in the antecedent of the conditional. There appears to be no more than a stylistic difference between (4) and either of the following:

(5)   If snow is, as I suppose, white, then grass is green.
      If, as I suppose, snow is white, then grass is green.

Instead of surrounding the parenthetical remark with commas, one may use parentheses, and in some cases one may instead use dashes.[19] Each sentence in (4) and (5) is radically different both semantically and pragmatically from

(6)   If I suppose that snow is white, then grass is green.

Unlike each example in (4) and (5), an utterance of (6) in a speech act need not engender assertoric commitment either to [[I suppose that snow is white]] or [[Snow is white]], and unlike each example in (4) and (5) the antecedent of (6) can be true without [[Snow is white]] being true. Further, consider the connective: '__, as I suppose', taking sentences as inputs and yielding sentences as outputs. (I shall identify this version with its variants.) When the placeholder is filled with a complement clause this connective can occur syntactically embedded arbitrarily deeply within conditionals, conjunctions, disjunctions, negation, modals, and attitudinatives. Pragmatically, this connective indicates that the speaker takes the complement to be true, and it does this regardless of how deeply

---

[19] The present approach will not attempt to discern semantic or pragmatic distinctions among these different devices. Further, I shall treat the first of the two sentences in (5) as having an LF corresponding to that of (4); in this respect I adopt an approach to transformationally relating these two sentences espoused by Emonds 1976, ch. 2.9. I shall refer to this as the Emonds Transformation.

embedded this connective is. Yet it does so by virtue of the meaning of the words of which it is composed, as is shown by the fact that a different although analogous role is played by such connectives as '__, as I presume,' '__, as I once held,' and '__, as many people think,'.

Our hypothesis for explaining the pragmatic features of the expression '__, as I suppose,' shall be that it is a robust, weak ifid in any utterance of which in a speech act the speaker shows her assertoric commitment to the semantic value of its complement. For first of all, for all illocutionary forces f, the sequence $< f, [[\ldots S_i \ldots, as_i \text{ I suppose}, \ldots]], \vdash [[S_i, \text{ I suppose}]]>$ is illocutionarily valid. What is more, $< \vdash [[S_i, \text{ I suppose}]], \vdash [[S_i]]>$ is also illocutionarily valid. It follows that

$$< f \ [[\ldots S_i \ldots, as_i \text{ I suppose}, \ldots]], f_\Delta \ [[S_i]]>$$

is illocutionarily valid, but this is just to say that '__, as I suppose,' is a weak ifid. Because it embeds arbitrarily deeply, it is also robust. The hypothesis that '__, as I suppose,' is a robust, weak ifid accounts for the fact that one who utters any of the sentences in (4) or (5) in a speech act shows her assertoric commitment to [[Snow is white]]. Analogous hypotheses suggest themselves for other first-person attitudinatives such as '__, as I presume,' and '__, as I conjecture,'. Likewise for speech-act verbs in the first person, for example, '__, as I assert,'. For brevity 'parenthetical attitudinatives' will in what follows be used to denote the class of parentheticals of the form, '__, as NP V_S,' where V_S is an attitude or speech act verb.[20]

Two parenthetical attitudinatives can occur in a coordinate position, and one can embed within another. In both cases the parenthetical attitudinative behaves in a manner our approach would predict. As an example of the first kind of case we have

(7)   If [snow is white]$_i$, [as]$_i$ John thinks, and [as]$_i$ Mary denies, then grass is green.

The thesis that '__, as John thinks,' is a weak indicator of assertoric commitment explains why one putting forth the displayed sentence in a

---

[20] The class of parenthetical attitudinatives as defined here is meant to exclude constructions of so-called free indirect style, such as are found in 'What a pleasant surprise, she giggled,' and 'I don't particularly, he rapped'. These do contain weak indicators of assertoric commitment. However, as observed by Reinhrdt 1975, parentheticals such as these do not embed arbitrarily deeply. For this reason such expressions are not robust.

speech act shows her assertoric commitment to [[John thinks that snow is white]]. Ignoring the complication presented by the connective 'and' in the second parenthetical, the thesis that '__, as Mary denies,' is a weak indicator of assertoric commitment explains why one putting forth (7) in a speech act shows her assertoric commitment to [[Mary denies that Snow is white]].

As an example of one parenthetical attitudinative nesting within another, we have,

(8)   If [snow is white]$_i$, [[as]$_i$ John, [as]$_{ii}$ we all know, thinks]$_{ii}$, then grass is green.

In light of the Emonds Transformation (see note 19), we take it that [[NP, as NP$\square$ V$_s$, VP]]=[[NP VP, as NP$\square$ V$_s$]], whence

f [[...If [snow is white]$_i$, [[as]$_i$ John, [as]$_{ii}$ we all know, thinks]$_{ii}$, then grass is green ...]].
$= f$ [[...If [snow is white]$_i$, [[as]$_i$ John thinks]$_{ii}$, [as]$_{ii}$ we all know, then grass is green ...]].

The hypothesis that '__, as John thinks,' is a weak indicator of assertoric commitment explains why a speaker uttering (8) in a speech act shows her assertoric commitment to [[John thinks that snow is white]]. The hypothesis that '__, as we all know,' is a weak indicator of assertoric commitment implies that a speaker uttering (8) in a speech act shows her assertoric commitment to

[[We all know that [as]$_i$ John thinks]].

But as we have observed, [[[as]$_i$ John thinks]]=[[John thinks that snow is white]], whence a speaker uttering (8) shows her assertoric commitment to

[[We all know that John thinks that snow is white]].

Parenthetical attitudinatives, then, possess semantic content, and indeed truth conditions, while also playing an expressive role. We thus cannot infer, from the premise that an expression has semantic content, to the conclusion that it can't also be a device of self-expression.

7
Expressive Qualities

In self-expression we show what's within, sometimes overtly, sometimes involuntarily, and at other times in ways that are between these two extremes. In certain cases we show what's within by showing *that* things are thus and so within us; in other cases we show what's within by making some aspect of what's within literally perceptible. I have been promising for many pages to explain and substantiate a third way in which we show what's within, namely by showing how some aspect of our experience feels. Now it's time to make good on that promise.

Evidence might show me that a friend is in a state of anguish without enabling me to know how she feels. Showing-that, that is, might justify my belief about how things are with someone without enabling me to know how that situation feels to her. The same goes for showing-α: Nothing in our account of part–whole perception, and the way in which we've invoked it to support the claim that emotions can be literally perceived, implies that a person perceiving an emotion will know how that emotion feels. For all we know, a Vulcan such as Mr. Spock can perceive an emotion on the face of a human being without having the faintest clue what that emotion feels like. Spock may be able to come to her aid in various ways but evidently will be unable to empathize with her.

We can enable others to perceive our experiential or emotional state without showing them how that state feels. Further, while when we show how something feels we in some cases demonstrate the presence of that state in us, we characteristically do more than this. In particular we characteristically enable others with the capacity for empathy to know what that state feels like. Such knowledge of how something feels is facilitated by affinities between sensations on the one hand, and emotions and moods on the other. Further, such knowledge is not limited to knowledge of emotion. It could also convey information about how one's experience feels, including even perceptual experience. Accordingly, the present chapter will enable

us to make progress on a question raised in Chapter 2 (under dictum 2.1.7) concerning the extent to which experiential states (pains, sensations, etc.) can be shown, and thus the extent to which they can be expressed.

Neither showing-that nor showing-α shows how someone's emotion or experience feels. Since inference from evidence and perception are the most intensely studied sources of knowledge in the cognitive sciences, one might wonder whether anything at all could show me how someone else's emotion or experience feels. What's more, if I try to find out by having that emotion or experience myself, I might remain in doubt as to whether what I feel is at all like what she or he feels. In this chapter I will first (7.1) explain that notion of showing more germane to skill than either propositional or perceptual knowledge. In 7.2 I then forward a theory that helps to account for this by hypothesizing a common three-dimensional space into which we can map both emotions and experiences. This puts us in a position, in 7.3, to account for a form of showing what's within that involves sharing the phenomenal character of an emotion or experience. That in turn helps us to understand how self-expression facilitates empathy, and I offer an account of this connection in 7.4. On that basis I also consider in 7.5 some theories of expression in the arts, focusing primarily on musical expression. I canvass some major theories of this kind and offer an alternative that I argue to be superior to all of them. I also offer an account in 7.6 of how representational art can be expressive without depicting anything that is itself expressive.

7.1. Showing how

One way to learn how to do something is to be shown how to do it. You might show me how to tie a shoe, how to find a constellation in the night sky, or how to calm down in a stressful situation.[1] In the first case you might give a visual demonstration of what to do with laces. In the second case you show me where to point the telescope and what configuration to look for. In the third case you might show me how to imagine a situation

---

[1] You might show me, and with luck I will learn. A teacher might show a class how to do integration but the students might be too bored or dense to grasp what she has taught. Accordingly, from the fact that A shows B how to do something, it does not follow that B thereby comes to know how to do it. In this respect showing how is like showing that.

that would be relaxing; perhaps I am to envision walking on cool green moss, or sitting on a beach with waves lapping my legs. If I do my work as a pupil, what I learn from teaching of these kinds will be skills, know how, rather than, or at least not just, knowledge that.

You teach me how to do something, and if I am doing my job as a student I will come to know how to do it. You might codify that lesson in a book or in some other artifact. A geometry book will show me how to prove a theorem. If I grasp what it teaches then I will know how to prove that theorem. A recipe book will show me how to cook a soufflé. Here too if I grasp what it teaches I will know how to cook a soufflé. Observe that in both cases, reading the book might not be enough for me to learn how; I may have to practice the theorem or soufflé a few times before I get it right. Know-how is often only possible with some practice.

Consistent with our pattern thus far, a painting might show me how a certain man's hair is colored, and if I have appropriate perceptual apparatus I will thereby learn how it is colored. That is only possible if the sitter's hair is so colored, and only if the painter and other aspects of the transmission of information are reliable. If all these things are in place, then I can learn how that man's hair is colored. If I retain this knowledge I will be able to discriminate this color from others. That is a skill that I might retain for a while and then lose, and when I lose that skill it will no longer be true that I know the color of the man's hair except, at best, propositionally ("It's russet"). Further, if I learn what his hair looks like, then I might be able to visualize that color in both his and his picture's absence. However, this ability to visualize is not a necessary condition of my knowing that color. (I know what sulfur smells like without its being the case that I can "olfactorily" image that smell, and I know what fuschia looks like without my being able to form a mental image of that hue.) We will see below that just as empathizing is something that we do rather than something that befalls us, to empathize with another it is not enough that we have had some experience of what they are going through. In addition to knowing what they are going through, we must "feel with them". The hard work comes in knowing how to elucidate the expression with which I've just shuddered.

A painting may not accurately and reliably portray the color of a man's hair. If it doesn't, then it doesn't convey knowledge of what his hair looks like. However, even in that case, it may still provide me with qualitative

knowledge, and thereby a skill, because it exemplifies the qualitative information that it also represents. It enables me to know what russet looks like because in looking at it I perceive that color. If I've never seen russet before, I learn something new. Further, even if I have seen that color before, this painting can activate that knowledge and thus provide me with the ability to visualize that color in its absence. So a painting can both provide knowledge and enable me to bring what I know into consciousness.

How distinct is knowledge how from knowledge that? One could try to conceptualize the aforementioned skills as knowledge that rather than knowledge how, suggesting, for instance, that knowing how to tie a shoe is just a matter of knowing that the way to tie a shoe is *this* (where one demonstrates the shoe-tying technique). Presumably here the shoe-tying demonstration is one that the knower must perform, rather than just pointing to someone else demonstrating the skill: A child does not demonstrate mastery of shoe-tying simply by pointing to an adult's demonstration of how it is done. But if this is correct, then knowledge-that in such cases presupposes a skill, precisely the skill exhibited in the aforementioned demonstration. For this reason, while knowledge-how may be analyzable in terms of propositional knowledge, that propositional knowledge itself presupposes a skill. Hence while knowledge-how may presuppose knowledge-that, it cannot be analyzed in terms of knowledge-that without remainder.[2]

Sensory experience typically gives us both knowledge-that and knowledge-how. Gazing at a mountain range I can learn how many peaks there are, what sort of vegetation they have, and whether there is evidence of recent fire. These are forms of knowledge-that. In addition, in gazing at the range I can acquire knowledge-how. For instance, I can learn what a particular combination of haze and vegetation looks like from a distance, or how a vast expanse of trackless wilderness looks. I learn how these things look; from other experiences I might learn how sulfur smells, how coconut tastes, or how a shark's denticles feel to the touch. In addition, knowledge

---

[2] Stanley and Williamson 2001 argue that all alleged cases of knowledge-how are analyzable as cases of knowledge-that, while the converse relation of analyzability does not hold. However, their "analysis" appeals to what they call "practical modes of presentation", and as Rosefeldt 2004 points out, in lieu of an elucidation of this notion we cannot tell whether the putative analysis succeeds without remainder.

of how something looks, sounds, tastes, and so on, seems to equip us with a skill. If I know how the red of an apple looks, or how the sour of a lemon tastes, I know how to recognize that color by its look, and that taste by its flavor. Although the connection is less reliable, I might also know how to imagine the red of an apple or the sour taste of a lemon.[3] These are skills lacked by someone born blind and lacking the sense of taste.

Not only might perception show me how something looks, sounds, and so on, and thereby equip me with skills; it might also provoke in me, and thereby give me knowledge of, emotions and moods. In watching a small child get hit by a car I feel horror; in listening to the surf I feel calm. An unfortunate turn of events could show me how desperation feels; a fortunate one could show me ecstasy or serenity. In all these cases I learn how an emotion or mood feels if I did not already know, and these experiences have shown me how that emotion or mood feels. What is more, even if I did know how that emotion feels but this knowledge was dormant, such experiences might activate that knowledge in me. In either case, subsequent to the experience I will if my memory serves me right be possessed of a skill: as before, a skill of being able to recognize that feeling by how it feels.

This phenomenon of knowing how a feeling feels will consist in different things for different people, and might even vary over time for a single person. For instance, one person might come to recognize the onset of her anxiety by a creeping sensation in her skin; another might recognize the onset of his anxiety by a slight sense of vertigo. Again, a person might recognize his trepidation by an unusual appearance in an object of perception, as in this passage from John Cheever in which a man perceives the effects of his attempt to abstain from tobacco and alcohol:

On Sunday I sneaked seven cigarettes in various hiding places and drank two Martinis in the downstairs coat closet. At breakfast on Monday my English muffin started up at me from the plate. I mean I saw a face there in the rough, toasted surface. The moment of recognition was fleeting, but it was deep, and I wondered who it had been. Was it a friend, an aunt, a sailor, a ski instructor, a bartender or a conductor on a train? The smile faded off the muffin, but it had been there for a second—the sense of a person, a life, a pure force of gentleness and censure, and I

---

[3] The connection is less reliable because I can know how the sour of a lemon tastes without being able to imagine that sour taste in its absence. The point comes out even more clearly with the sense of smell. I know how sulfur smells, but may not be able to imagine smelling sulfur.

am convinced that the muffin had contained the presence of some spirit. As you can see, I was nervous. (Cheever 1960)

His seeming to perceive the face in the muffin shows the character (and us) how his anxiety feels or at least some aspect of that anxiety. On the other hand, not only would I not expect my anxiety to feel this way for me, it is also unlikely that this character's anxiety would feel the same way for him on another occasion. In contrast to the smell of sulfur or the taste of coconut, how a mood or emotion feels can take quite different forms for different people, and can change over time for one person.

As with the case of knowledge of how perceptual experiences feel, experiences of moods and emotions might also equip me with an ability to imagine how an emotion or mood feels; however, as before, the connection is a relatively loose one. Unless I am in a situation that elicits feelings of desperation or anxiety, I may have difficulty imagining that feeling. This is why some psychotherapists have in recent years begun using virtual reality devices to help patients face their phobias: those devices provide patients with experiences that they cannot call up through imagination or memory on their own, but that nevertheless enable them to face the fears that those experiences elicit.[4]

Just as what qualitative knowledge we acquire from experience depends upon our sensory capacities, so too what emotional knowledge we acquire from experience depends upon our physiological, as well as ambient emotional makeup. A child may know no lust even upon seeing what provokes lust in an adult. A sociopathic adult will feel no horror even as the vehicle smashes into the child. Perhaps, as we have suggested, Mr. Spock is capable of no emotion at all while having the same perceptual experience as human beings do. Less fancifully, different individual histories will result in one and the same experience triggering different affective responses in viewers: a certain smell will trigger a fond memory of a childhood event in me but not in you. Even taking into account these sorts of variation, we may still say that perception often activates emotion, and can thereby provide or activate knowledge of how an emotion feels.

Self-expression requires showing one's introspectible state. In addition to our account of making that state literally perceptible, or giving a

---

[4] See Sam Lubell, 'On the Therapist's Couch, a Jolt of Virtual Reality', *The New York Times*, 19 February 2004, section G, page 5, column 1.

demonstration of the presence of that state, we are now ready to understand a third form, namely self-expression as showing how one feels. As the locution 'showing how one feels' is used in English, I can show how I feel simply by grimacing, or by saying that I am annoyed. In some cases, however, I wish to enable others to know how it feels to be experiencing what I am; in such a case I want others to know how my melancholy, or exuberance, or anguish, or sense of loss feels. How might I do that? Here are some ways:

1. I might get you to experience whatever it is that put me into the affective state I am in, or at least something else sufficiently similar so that it is likely to produce similar effects on you. For instance, I send you the same cloying salesman who has raised my ire.

2. I might describe my feeling and ask you to imagine feeling that way. For instance I describe my debilitating melancholy in the hopes of getting you to imagine how I feel. To this end I might describe, as John Cheever's character does, how an object appears to me.

3. I might draw your attention to something external that in some way corresponds to the way I feel. For instance, I point out the weeping willow as corresponding to how I feel by looking the way I am inclined to behave, namely droopy and weepy. Or I point out the raging storm and remark, ''That's how I feel.''

Each of these procedures is at best a fallible means of getting you to understand how what I am experiencing feels. In cases of the first kind I might remain in doubt whether my own response to the salesman is much like yours. In cases of the second kind I might doubt whether my imagination has replicated your feeling. In cases of the third kind I might be unsure in just what way the raging storm is like your feeling. Further, just as there are familiar quandaries about whether I can know what your experience of yellow or taste of orange are like, so too we might wonder whether I can ever know what your experiences of melancholy, exuberance, or anger are like. In Section 7.2 we will, however, see interpersonal, as well as intrapersonal, limits on the possible forms that, say, feelings of melancholy, anger, or surprise can take. In light of these limits we then (Section 7.3) will be able to see that showing how one feels can enable others to know what your experience is like well enough

to provide them with what they characteristically seek from this form of self-expression, namely empathy.

We are also in a position to understand how one might convey knowledge of how a certain emotion, mood, or experience feels without its needing to be an emotion, mood, or experience being felt by anyone. Here, then, is one area in which expression conveys knowledge that goes beyond what is within a certain individual, and has the capacity instead to provide knowledge of a more universal kind. As we will see, showing how something feels is deeply bound up with artistic practice. This is the main topic of Sections 7.4 and 7.5.

7.2. Congruence of sensation and affect

Experience provides propositional and qualitative knowledge, and can be the catalyst for my knowledge of emotions. As the source of these latter two, experience provides me with knowledge how rather than knowledge that; it shows me how things look, sound, and so on, or how certain emotions feel. In addition, experiences themselves often have qualities enabling us to convey knowledge in indirect ways. To see how this is so, notice that many experiences are painful, pleasant, sour, bitter, sweet, or soothing. To someone with a certain physiology, a sound pattern will be soothing, a smell will be unpleasant. Again, to someone with a certain physiology, a chord will have a melancholy sound while a color might seem exuberant. (To creatures differently endowed, experience with these objects might not have anything like the same affective dimension, if any affective dimension at all.) Now, to say that an aroma is acrid to certain creatures is not to say that any such creature experiencing that aroma will enter into a sad or other negative emotion or mood. Pain is also inherently unpleasant, although of course to a masochist that displeasure might give rise to pleasure as well. The unpleasantness of pain might make a person experiencing it unhappy, but need not do so. So too an acrid smell such as the combination of sweat and rust in an old and heavily used bus is unpleasant even if it triggers appealingly bittersweet memories in me.

These points are germane to the well-documented phenomenon of cross-modal congruence, in which some sensations within one sensory modality seem to bear more of an affinity to some sensations within

another sensory modality than to others. Intuitively, we think that yellow is more like the sound of a piccolo than it is like the sound of an oboe; that the smell of sulfur is more like rough than it is like smooth; that the taste of lemon is more like the minor chord C−E flat−G than it is like the major chord C−E−G, and so forth.[5] Many such phenomena are borne out experimentally. To take two examples among many: subjects reliably match the louder of two tones with the brighter of two spots of white light, as well as matching higher-pitched tones with brighter lights (Marks, Hammeal, and Bornstein 1987).

More generally, our emotions and moods may be described along a number of dimensions, including the following three:

intense/mild
pleasant/unpleasant
dynamic/static.[6]

Anger is intense, slightly unpleasant, and highly dynamic. I take it that neither the intensity nor the unpleasantness of such an emotion is in need of elucidation. What does it mean to say that anger is dynamic? Recall that in Chapters 3 and 4, in the course of discussing the ''affect program'' conception of basic emotions, we suggested that emotions tend to have a life of their own; one in the grip of anger will be disposed to actions—raising her voice, kicking animals and furniture—that are liable to occur unless she makes an effort to prevent them. If she does not make such an effort but could do so, these acts are things she allows rather than does. (Observe, however, that in extreme cases she may be unable to prevent such actions.) Likewise, fear makes us flee unless we take steps to prevent flight (imagine being rushed by an angry dog that had given no warning bark); disgust makes us retch unless we curb the impulse to do so (imagine finding a large slug on your arm). These three emotions, then, are relatively dynamic. By contrast, sadness tends to produce the cessation of action and so is static, as well as being unpleasant and moderately intense.

---

[5] These intuitive judgments are borne out by a variety of experimental investigations surveyed in Marks 1978, and further developed in Marks 1987, 1995.

[6] These three dimensions are close to those proposed by Hartshorne 1934. They are not beyond dispute, and others have been proposed. For a survey of options see Marks 1978, chapter 3. Indeed, Marks 1995 refers to a ''constellation of multidimensional relations connecting perception in different sense modalities'', (p. 213). Our approach does not depend on the three dimensions mentioned in the text being precisely the correct ones.

Not only can emotions be characterized in these crude multi-dimensional terms; sensations can be so characterized as well. Leaving aside their representational characteristics (for instance an experience of an unpleasant object will itself often be unpleasant), sensations can often be characterized as intense/mild, pleasant/unpleasant, dynamic/static:

Vision. Some visual patterns, for example, seem to scintillate while others seem still; this is due to facts of the human visual system that are the proper topic of a psychology and physiology of perception. Scintillating patterns are dynamic. Again, yellow, orange, and red are more intense than blue and green, for instance, and yellow seems more dynamic than orange; both are more dynamic than either blue or green (Wilson 1966). It is not clear that non-representational aspects of visual experiences differ significantly from one another along the pleasant/unpleasant dimension.

Audition. Some chords, for instance the C major, feel at rest and thus static while others feel dynamic, and thus in need of resolution (Wicker 1968; Kivy 2002). Some sounds are screeching or eerie, and are unpleasant, while others are smooth or sweet, and are pleasant. Sounds are of course felt to vary in intensity with volume; but they are also felt to vary in intensity with pitch (Marks 1995).

Taste, smell, and touch. Clearly taste and smell sensations differ along the pleasant/unpleasant dimension—what some researchers refer to as "hedonic tone". Jeddi 1970, for instance, finds cross-cultural evidence that the sensation of warmth is more pleasant than is the sensation of cold. Taste and smell also differ along the intense/mild dimension. It is not clear that they differ in any significant degree along the dynamic/static dimension. Likewise for tactile sensations.

Our sensory modalities, then, enable us to make intermodal comparisons. Here is an empirical hypothesis to explain the basis of such judgments: Normal members of our species have epistemic, generally non-conscious, access to a three-dimensional coordinate system in which elements of a sensory modality can be mapped; the same goes for the qualitative components of such moods and emotions as have qualitative components. For each sensory modality, there will be a set of elements of that modality: The set of all visual experiences I have or have had,

the set of all olfactory experiences I have or have had, and so forth. For simplicity, let's just refer to each such set as a modality itself—so that the visual modality will be a set of visual experiences, and so forth. Members of each sensory modality, as well as each emotion or mood with a qualitative character, may now be mapped onto a point or set of points in the three-space determined by the dimensions given above. An experience with a distinctive degree of intensity/mildness, pleasant-ness/unpleasantness, and dynamism/stasis will be mapped onto one point in the three-space. By contrast, an experience that is, say, pleasant and dynamic, but neither intense nor mild, will be mapped onto a set of points rather than just one. Experiences with a distinctive character in only one dimension will be mapped onto a plane within our hypothesized three-space.

My hypothesis, then, is that in having experiences we also place those experiences onto distinctive points or spaces in the aforementioned three-space. I also propose that we are able to discern the proximity not only of two thus-placed experiences within the same modality, but also of two experiences from different modalities. Suppose that visual experi-ence V has a degree of dynamism d, and that auditory experience A has that same degree of dynamism. If someone asks what my expe-rience A is like, or how it feels, I can provide a partial answer by making visual experience V available to him—most likely by providing a visible object. If the questioner manages to have that visual expe-rience, she may then come to know what A is like. This notion of "what it's like" is, however, notoriously vague. More precisely she will come to know how intense A is, and this is a substantial piece of knowledge about A. The same point applies to different modalities of assessment.

This account of what underlies our judgment of "congruences" or affinities among elements of different sense modalities, or elements of a sense modality to one or more emotion, does not imply that discerning such a similarity requires a conscious process of judgment and/or comparison. Instead, the process of discerning such similarities falls to the so-called adaptive unconscious, which was adumbrated in Section 4.3, and which has been at the center of interest in much recent experimental psychology. According to this research, the adaptive unconscious is responsible for a great deal of our "automatic" behavior, including such things as judgments

about a perceived object's distance from us and its relative location, syntactic and semantic processing, and many of our impressions about other people.[7] Employing the doctrine of the adaptive unconscious, we may say that sensitivity to intermodal congruence need be no more conscious than is our judgment of a perceived object's distance; or the process of interpreting a sentence newly encountered. Unlike the processes that result in our judgments about how far away an object is, and like semantic processing, our judgments of intermodal congruence may in certain cases be made conscious with sufficient acute introspection. Rather than just, say, intuiting an affinity between elements of two different modalities, we might in certain cases become introspectively aware of the basis of that affinity, for instance, in order to express ourselves. We may, however, find evidence for the awareness of intermodal congruences without yet knowing whether any such awareness can be the subject of introspection.

7.3 Showing what's within, part iii

Our sensitivity, be it conscious or unconscious, to intermodal congruences helps explain how we are able to convey information, if only of a limited kind, about aspects of our experience or affect. If you've never heard a piccolo, by telling you it is like yellow I give you some know-how, namely partial knowledge of how that instrument sounds, and thereby some ability, albeit limited, to recognize it by its sound. I might also enable you to imagine how it sounds, but, as before, this connection is not entirely reliable. In any case, these intermodal comparisons enable us to show some aspects of how our experiences feel to others who do not know how such experiences feel or in whom knowledge of how they feel is dormant.

---

[7] See Wilson 2002 for an overview of the cognitive unconscious as it figures into recent cognitive and social psychology. I differ from Wilson in one important respect. Wilson defines the cognitive unconscious as an area of mental processing inaccessible to conscious awareness. It is thus distinguished from the traditional notion of the ''preconscious'', an area of consciousness outside conscious awareness but accessible to it. Although the change does not drastically affect his theory, I would hold that many of the phenomena that he treats as inaccessible to consciousness are in fact accessible to it. Wilson later in this work in fact rescinds his overall policy for some emotions, and I would go a step farther and treat the cognitive unconscious as on the whole preconscious, with exceptions made for specific cases that are not open to introspection.

Not only are intermodal comparisons possible, elements from a given modality can be compared with emotions and moods. As we have seen, anger is intense, slightly unpleasant, and dynamic. Sadness is intense, unpleasant, and static, although sadness veering on anguish is intense, unpleasant, and dynamic. Disgust is intense, highly unpleasant, and dynamic but less dynamic than anger or anguish. These characterizations enable us to compare elements of one sensory modality with an emotion or mood. The major triad C−E−G is congruent, according to this system of measurement, with confidence or cheerfulness, for both are intense, pleasant, and relatively static. The color yellow is congruent with exuberance, for both are intense, pleasant, and dynamic.

We have, then, intermodal conguence, as well as congruence between elements of a given sensory modality and our emotions and moods. If congruence of the former sort exists, and we are aware of it when it does, then we can make sense of how I can show you, along one or more dimensions, what my experience is like by presenting sensory stimulation containing elements congruent to that experience. I thereby enable you to know certain aspects of how what I am experiencing feels. Similarly, if congruence of the latter sort exists, then we can make sense of how I can show you, again along one or more dimensions, what my emotion is like by presenting sensory stimulation containing elements congruent to that emotion. When I do that, I enable you to come to know aspects of how that emotion feels, rather than enabling you to perceive that emotion. Will these limits prevent my showing you enough of how I feel to make our combined efforts worthwhile? As we'll see in the next section, one standard of success is whether the audience is able to empathize with me enough to render aid, form an alliance, mate, or such like.

Return to our example from Chapter 2 in which Nathaniel Hawthorne describes Hester Prynne emerging from prison:

Those who had before known her, and had expected to behold her dimmed and obscured by a disastrous cloud, were astonished, and even startled, to perceive how her beauty shone out, and made a halo of the misfortune and ignominy in which she was enveloped. It may be true, that, to a sensitive observer, there was something exquisitely painful in it. Her attire, which, indeed, she had wrought for the occasion, in prison, and had modeled much after her own fancy, seemed to express the attitude of her spirit, the desperate recklessness of her mood, by its wild and picturesque peculiarity.

The "wild and picturesque peculiarity" of Hester's attire doesn't make her desperate recklessness perceptible—that recklessness doesn't seem to have visible characteristic components, nor does it appear to be perceptible in other ways. Furthermore, her attire might show, or help to show, that she is feeling desperately reckless, but it does more than this. It enables us to know how she feels because experience of that attire is congruent, in a way we now know how to articulate, with the experience of desperate recklessness. Hence the exquisite pain for the sensitive.

All these attempts to show others what our emotions or moods feel like might seem to run afoul of the, or at least one version of the, philosophical problem of other minds. Is not the "qualitative feel" of another's experience something to which I in principle cannot have access? Likewise even for my own experiences considered diachronically: could not things appear different to me now from the way they did yesterday without my being any the wiser?

Notice, first of all, that some kinds of inter- as well as intrapersonal inversions do not seem possible. It is difficult, for instance, to see how there could be an interpersonal inversion as between pain and pleasure. This would require that the experience that I feel upon cutting my hand with a knife is like the experience you feel upon stroking velvet. What could that mean? Unless your nervous system is detectably out of order, your cutting your hand must, like mine, produce an unpleasant experience. Whether your experiences resulting from bodily damage are exactly the same as mine is not at issue; the point rather is that both experiences must be unpleasant. Likewise, while I may be unsure whether the lemon looks to you just as it does to me, even when we are both in normal lighting conditions and our senses are in good working order, I will feel sure that however it looks to you, its color will look exuberant. So too, while aware of the possibility of an interpersonal inversion of sound qualities, I can nevertheless rule out the possibility that a minor chord sounds sad to me but happy to you. We can, that is, know something non-negligible about how one another's experiences feel, even if that knowledge still leaves some residual room for undetectable variation.

It might also be suggested that if the use of intermodal congruence could really enable others to know how my experience feels, then it would, *per impossible*, solve what-it's-like-to-be-a-bat problems as well. Could an intelligent bat-like creature possessed of echolocation, or shark-like

creature possessed of electroreception, show me how its experience feels by finding experiences to which I have access that are congruent to theirs along the three dimensions discussed above? I suggest that they might very well do this, and if they did so they would give me some idea of what their experience is like. It is an empirical question whether any of the shark's experiences vary along the dimensions of pleasant/unpleasant, intense/mild, or dynamic/static. However, if any of them do, we can learn something of how the shark's electroreceptive experiences feel. I do not claim that the phenomenon of intermodal congruence answers all questions about how the world seems to a creature with a different sensory modality from any that we own. I do, however, suggest that doctrines holding that such experiences are ''fundamentally alien'' to us are overblown; their plausibility in the very least depends on whether these forms of experience can be characterized in such a way as to allow us to discern intermodal congruence.

We now have a third means of self-expression in addition to the two elucidated in earlier chapters, expression-as-making-perceptible and expression-as-demonstrating. Exploiting a perceived congruence between sensory quality S and emotion, experience, or mood E, I can present you with S with an intention of showing you how E feels. Unlike expression-as-making-perceptible and expression-as-demonstrating, this form of showing what is within puts you in a position to know how E feels—or more precisely how ϕ E feels, for one or more values of a ϕ that ranges across the three dimensions we've mooted.

Further, while it's clear that one can express both cognitive and affective states, we have left open since Chapter 2 whether one can express experiences such as a pain, a smell of vanilla, or the taste of lemon: While challenging the contention of some authors that doing so seems a conceptual impossibility, we suggested that expression of experience might require ingenuity of the sort we find among great artists. Now we can explain the sort of ingenuity required. If I can show you how my experience feels along one or more of the three dimensions under consideration here (intense/mild, pleasant/unpleasant, dynamic/static), that will give you knowledge of how ϕ that experience is, where ϕ ranges over one or more of those dimensions. I could show you something of how my taste of vanilla feels by playing a smooth, sweet chord on the piano. I could show you something of how yellow looks by

playing the piccolo, or to take Locke's case, of how scarlet looks by playing the trumpet. Intermodal congruence, then, enables us to show aspects of how our experiences feel, and thus enables us to express those experiences.

In earlier chapters we stressed how showing is a stronger relation than mere indication; the former, but not the latter, confers justification of a sort appropriate for knowledge and so is intimately bound up with knowledge. Showing-that makes knowledge-that available; showing-α makes perceptual knowledge available, and showing-how makes available knowledge of how an emotion or experience feels. The norms governing speech acts help to make the first kind of knowledge possible by serving as handicaps; what vouchsafes attempts to provide knowledge by showing how an emotion or experience feels? There are two questions here. First of all, we may ask how one shows (rather than just indicates) how one's own emotion or experience feels. Secondly, we may ask how one shows (rather than just indicates) how an emotion or experience feels whether or not it is one's own.

There may be no general answer to the first question; how justification is made available may depend on the case in question. For instance, a bit of behavior that shows how an emotion feels may depend for its credibility on an implicit assertion: My emotion feels like *this*—and then the agent demonstrates an object with phenomenal qualities she alleges to be congruent with her emotion. The assertoric, and thus handicap-involving backdrop of this act then accounts for such credibility as it has. By contrast, it is plausible that many of the automatic aspects of our expressive behavior show what they do by being indices of the qualitative characteristics of the emotions that cause them: I don't choose the register of my voice when I scream in pain, yet that register is surely congruent with an intensity that helps to characterize my pain.

How about showing an emotion or experience that is not mine, thereby creating a work that is expressive but not a self-expression? I suggest that in such a case the artist shows (rather than merely indicates) how that emotion or experience feels only when she has achieved a credibility within her community: she has a reputation for showing how emotions and other experiences really feel—as borne out by frequent experience with her artworks in which we find them correct to the emotion or experience in question. This may help account for the importance we

attach to authenticity among artists in spite of not demanding that they express emotions or experiences of which they are currently possessed.

An artist's credibility also depends on her ability consistently to put us in a position to empathize with others. However, let me repeat a caveat before we consider the role of empathy in expressiveness and self-expression. I have postulated the three dimensions of intense/mild, pleasant/unpleasant, and dynamic/static as the basis of intermodal congruence, as well as the basis of congruence between sensations and affect. That particular hypothesis may turn out incorrect while the more essential features of the present approach stand. Intermodal and sensory-affective congruences may depend upon other dimensions than those I have hypothesized: perhaps there are other dimensions beyond these, and a better theory might not even include all these three. The robustness of the phenomena of intermodal and sensory-affective congruence strongly suggests, however, that some such dimension or dimensions must exist to make these congruences possible.

7.4. Empathy

What is the value of enabling you to know how my emotion or experience feels, and is that value, if such there be, relevant to self-expression? After all, in light of what we've learned in earlier chapters, I could show you my anger by making it perceptible, or instead by demonstrating its presence with compelling evidence. Why would one need a third way beyond these two? My answer, in broadest outline, is that one of the signal virtues of showing how my emotion or experience feels is that doing so puts others in a position to empathize with me. We empathize with others when we imagine how they feel, but where the imagination in question must with reasonable accuracy capture that feeling. Thus, while imagination is not essentially constrained to accuracy (I can imagine winged horses and golden mountains), my imagination "puts me in your shoes" only if what I imagine is, or is relevantly congruent with, what you're feeling. Once in your shoes I will then be more likely to come to your aid, assuage your pain, become an ally, and so forth.

The imagination required in empathy is thus constrained by a requirement of "direction of fit": it must track how things are with you if it's to count as empathizing with, say, your shame rather than your regret.

It is also constrained by the requirement that I imagine *myself* feeling the feeling that I am ascribing to you; it is not enough that I imagine your, or for that matter someone's, feeling what I am ascribing to you. Call these the *fidelity* and *de se* requirements, respectively. These two requirements help explain why it can be so challenging to empathize with those from radically different cultures or with views deeply opposed to our own. For instance, I personally would have trouble empathizing with Aztec priests, or with someone who bombs an abortion clinic. I have difficulty imagining myself feeling compelled to remove a beating heart from a live person, or feeling the necessity of killing anyone involved in abortions. Even if I learn what these people feel, I have trouble imagining myself feeling like that.

The difficulty of meeting the fidelity and *de se* requirements for empathy also helps to explain why it is a considerable achievement when an author, screenwriter, documentarist, or photographer makes such empathy possible. For example, in the famous anti-war documentary, *Hearts and Minds*, we cut from a shot of General Westmoreland telling us that Asians don't have the same concern for the sanctity of an individual's life as is common in the West, to a shot of a Vietnamese mother weeping uncontrollably at the grave of what is presumably her lost son. Our empathy with her grief requires our imagining ourselves losing a child to war, and thereby gives us a glimpse of her suffering. In so doing, we see the absurdity of Westmoreland's pronouncement.

Again, a novelist might enable me to imagine being ostracized within a small rural community. Jane Hamilton does this in her *A Map of the World*. In its skillful depictions of various conversations, tones of voice, facial expressions, and so on of characters, Hamilton shows us what a sense of social isolation feels like. She does this by enabling me to imagine being in that situation, and thereby enables me to know how I would feel if I were in that situation. Even if I have never felt ostracized, the excellence of the novel consists at least in part in its ability to show me how that would feel.[8]

By showing you how an emotion, mood, or experience of mine feels, I might, if you are appropriately attentive, equip you with two distinct skills: first, the skill of being able to recognize that emotion, mood, or

---

[8] I elaborate on these points in Green forthcoming b.

experience by how it feels; secondly, the ability to imagine how that emotion, mood, or experience feels. Because of the latter, by showing you how an emotion, and so on, feels, I might enable you to imagine yourself feeling what I am feeling. In doing so you can fulfill the faithfulness and *de se* requirements. For this reason, by showing you how an emotion, mood, or experience of mine feels I might enable you to empathize with me. The facilitating of empathy is, then, one of the things to which expression-as-showing-how-it-feels is suited, and distinguishes it from both expression-as-perception-enabling and expression-as-demonstrating.

I said that showing how an emotion, mood or experience feels might enable you to empathize with me. You might try but fail to do so, for reasons I'll return to in a moment. By contrast, many writers hold, for A to empathize with B, or more precisely for A to empathize with B's φ, where φ is an emotion, experience, or perhaps even a thought, A must feel φ and on that basis imagine being in B's shoes.[9] On this view, to empathize with your terror I must feel terror myself, and to empathize with your resentment of God I must resent God too. Of course this is not a sufficient condition; I don't empathize with your aching just-stubbed toe by stubbing mine. In addition to sharing your emotion (or experience), on this account I must use my own replication of your situation to imagine my way into what you are feeling.

Not only is this view intuitively implausible, writers seem to be driven to it by impoverished conceptions of the imagination and of emotion. It is of course *not* enough for me to empathize with your feeling of being ostracized that I have been ostracized in the past. Nor is it enough that I have been ostracized in the past and I am capable of calling up that memory into consciousness. One is not empathetic simply by virtue of having dormant skills. Instead I have to do something that makes me count as feeling with you. But it is a mistake to infer that this feeling with others requires actually duplicating their feelings in myself. After all, to get myself to feel ostracized I'd have to induce certain beliefs in myself, such as that I am being excluded from a group on the basis of inappropriate considerations. (I am certainly not endorsing the view that emotions are a

---

[9] See for instance Gaut 1999 and Plantinga 1999. Frith 1989 also takes empathy to require an actual sharing of emotions, writing, "Empathy presupposes, amongst other things, a recognition of different mental states. It also presupposes that one goes beyond the recognition of difference to adopt the other person's frame of mind with all the consequences of emotional reactions." (p. 144–5)

species of judgment; I am assuming something much weaker, namely that certain emotions require judgments or beliefs as necessary conditions.) But suppose I don't feel that way at this point; as it happens, I feel more or less accepted by the groups I care about. Writers like Berys Gaut (1999), Alex Neill (1996), and Ute Frith (1989) would infer that I am incapable of empathizing with your sense of ostracism, but surely that is untrue. It would be awfully nice of me to follow a Pascal-style routine to get myself to believe that I am being excluded, but by the time I carry this off it will probably be too late for my empathy to be worth anything to you. Instead, I could save a lot of time and effort simply by calling up into conscious awareness my *memory* of how I felt when I was ostracized in the past. On the basis of that conscious awareness, I now know how you feel, not dispositionally but occurrently. If I now go on to use this conscious awareness as a prop in which I imagine that you are feeling *this*, then I have empathized with you.

Gaut didn't claim that for me to empathize with your anguish it is sufficient that I feel anguish. He takes this as a necessary, but not a sufficient, condition. Following Neill 1996, Gaut also holds that, "empathy requires one imaginatively to enter into a character's mind and feel with him because of one's imagining of his situation" (1999, p. 206). So on this view, to empathize with another's feeling of φ, I not only have to feel φ, I must also use that feeling of φ as a kind of prop on the basis of which to imagine being in your situation. This further condition of imaginative identification seems eminently plausible. What it does not do is mandate any requirement that I actually feel what I imagine you to be feeling. Rather, it is enough that I be able to call into consciousness my experience of that feeling without actually reliving it, and then on that basis imagine my way into your situation. That seems to be enough to enable me to feel with you. For emotions and moods having a qualitative dimension, we now see that it is also enough to learn how they feel by finding their location in the three-dimensional space we suggested in the last section.

This in turn, however, raises the question why this elicitation of empathy should be of value. Some people are less empathetic than others; in certain forms of autism the capacity for empathy is considerably impaired. However, when I do elicit another's empathy, that is one large step in winning their aid. One who empathizes with my pain, fear, or sadness stands

a good chance of providing protection, of helping to nurse wounds, or of offering encouragement.[10] These claims hold at best for the most part and admit of many exceptions. However, it does seem in general that one who empathizes with my pain (fear, sadness) is more likely to come to my aid than if she were merely aware, by means of either perception or demonstration, of that pain (fear, sadness). This aid-eliciting dimension of expression-as-showing-how-it-feels is one important feature distinguishing it from expression-as-making-perceptible, as well as expression-as-demonstrating. Neither of these two latter phenomena is particularly associated with the elicitation of succor; expression-as-showing-how-it-feels is made to order, and is one reason why showing-as-ability-enabling plays a distinctive role in self-expression.

Attempts by self-expression to facilitate empathy will not be equally effective for all people. As we have mentioned before, one characteristic of certain forms of autism is the presence of high-level cognitive capacities combined with the absence of an ability to empathize with others. For this reason an autistic person might be left cold by watching a mother grieving over her dead child, while many others will be moved powerfully. Further, common empathic responses might be invoked for the sake of manipulation, as is common in advertising. In addition, an artist might invoke such responses in order to undermine some of our own beliefs about ourselves. For instance in Leon Golub's *Interrogation* series we often meet the smiling eyes of one of the ''interrogators'', while the victim's face is covered by a mask or a hand. As a result a first reaction is to empathize with the situation of the interrogator, perhaps to feel the challenge of extracting information from a diffident suspect, and maybe even the pleasure of having another person completely under our own power. We might do this half-consciously, before realizing that we would do better to consider the plight of the man bound in the chair. As we reflect on this sequence of reactions, we may feel disturbed at having identified with a perpetrator before identifying with his victim, and even further disturbed about having to realize that this reaction was both immediate and not fully conscious. Does each of us harbor a torturer within?

---

[10] Williams 2003 stresses these points in her evolutionary account of the facial expression of pain. Green 2003a offers a refinement of her account of the communicative role of such facial expressions.

7.5. Artistic expression

A central question for philosophical aesthetics asks what it might mean for an inanimate object such as a piece of music, a sculpture, a painting, or a poem to be sad or have some other affective property such as exuberance, anguish, edginess, or serenity. In fact this question arises not just for works of art or, for that matter, just for artifacts. We might also wonder how the wind can have a melancholy sound, or how the flowering tree manages to look exuberant. In addition, many things other than emotions can be expressed, so we do well to consider how it is that a bonsai might have a pensive look, or how a building might have a menacing demeanor.

The problem, then, posed at the right level of generality, is how a non-sentient object can have an affective or other expressible property. Many philosophers concerned with this problem point out that non-sentient objects are not capable of feeling such emotions as sadness, anger, or fear, and conclude that a puzzle arises from the fact that we ascribe qualities like sadness and anger to these non-sentient objects. From what we have learned in foregoing chapters, an emotion such as sadness, anger, or fear is a complex phenomenon only a component of which is its qualitative dimension. For all we have established thus far, then, anger might be an emotion with, say, four possible criteria (a qualitative feeling of anger, impulses to aggressive behavior including facial expressions, a judgment of a certain sort, and physiological changes such as increased blood flow to the arms) any three of which are sufficient for its attribution. In that case one could be angry without feeling angry, whence the fact that an inanimate object is insensate does not yet show that it cannot be angry.

Of course, many inanimate objects to which we ascribe emotional qualities also lack a physiology and dispositions to behavior as well. For this reason they would seem to meet either none, or too few, of the other criteria needed for ascription of affective states. On what basis, then, can we be ascribing such qualities as melancholy to the seascape, the yew tree, or the sonata? It might be suggested that we simply do talk this way, and that this way of talking, being one of the language-games we play, is not in need of any grounding. In response, we may agree that we do indeed talk this way but remark also that we should not accept without argument that it is *simply* the case that we do so. Perhaps we do so for a reason.

Moreover, while the fact that we talk this way may not be *in need* of any grounding, it might be illuminating to investigate why we talk the way we do. Otherwise it will be an open possibility that the attribution of sadness both to the weeping child and to the sonata is a play on words, just as it is a play on words to remark that I have two trunks in the driveway, one in my car and the other possessed by my pet elephant.

I don't assume, then, that the very intelligibility of ascriptions of affective qualities to inanimate objects has to be vouchsafed by a philosophical explanation. What I'll suppose is much weaker, namely that we *may* shed light on such ascriptions by investigating their basis. In this section I'll consider only the way in which expressiveness may be found in objects that are not representational, as well as in objects whose representational character has nothing to do with their expressiveness. (In the next section, 7.6, we will consider how expressiveness can be achieved through representation.) Note also that while philosophical aesthetics is concerned almost exclusively with emotional expressiveness, our discussion thus far suggests that this restriction is not mandatory. It is no less important to understand what it might mean for an inanimate object to express doubt, certainty, or indifference; likewise for items of perceptual experience.

One more methodological remark: it may be that expressiveness takes quite different forms and admits of quite different explanations in the various arts. However, I shall assume that all else being equal, the more we can explain the varieties of expressiveness in terms of a single principle or set of principles, the more powerful that explanation will be. The majority of theories of expressiveness in non-representational arts are primarily concerned with music. Nevertheless, an explanation that purports only to account, say, for how musical expressiveness works is liable to be trumped by another account that does at least as well as it does for music but that also applies to other art forms.

What, then, is it that makes a (non-programmatic) piece of music melancholy; in virtue of what is a color or pattern of colors and/or shapes exuberant? More generally, *in virtue of what does a non-sentient object A possess affective or experiential quality E*? Answers to this question fall into two categories, namely cognitive and non-cognitive theories.

Cognitive theories. According to theories of this kind, an object possesses an affective or experiential quality E just in case E can properly be discerned or

imagined to inhere in A by appropriately placed viewers (listeners, etc.) of A. That discernment or imagination might be due to the perceived resemblance of E to some other object (such as a sentient agent expressing her emotion) (Kivy 2002; Davies 1994), to an ascription of a state of mind to the imagined utterer of the work or some of its parts that best accounts for its distinctive features (Vermazen 1986), to being a best hypothesis on the part of an ideal listener as to what state of mind the composer of the music intended such a listener to hear in the passage (Stecker 2001), to being an ascription of a state of mind to the imaginary protagonist of the passage that figures in the best interpretation of the work taken as a whole (Robinson 1998), or to being an ascription of properties to an imagined persona or personae in the work (Levinson 1990, 1996, 2002, 2006; Maus 1997).

Non-Cognitive theories. According to theories of this kind, an object A possesses affective or experiential quality E just in case A is appropriately related to an emotional or other non-cognitive experience of E on the part of appropriately placed viewers (listeners, etc.). The relation in question might be a matter of causation as between the artifact and some affective state of the viewer (the arousal theory), causation as between an affective state of the creator of the artifact and the artifact, causation as between the creator's affect and those of the viewer (Tolstoy 1989), or of a dispositional relation between work and viewer (Matravers 1998, 2003). These non-cognitive theories do not require a belief, judgment, or imagining on the part of the viewer that the artifact in question possesses any affective quality.

The approach that falls naturally out of our investigations in this and previous chapters is what we might call the Expressiveness as Showing Theory. With apologies to the New Age semi-cult popular in the seventies, let us abbreviate this to 'EST'. We've said that all expression is a matter of showing what is within, and that this phenomenon in turn can take one of three forms—making perceptible, demonstrating, and showing how something feels. These aspects of showing naturally suggest an account not only of expressiveness as showing one's feeling (thought, experience, etc.), but also of expressiveness as showing *a* feeling (thought, experience, etc.) that need not be being felt. A painting might present, and thereby show, rage without it being a rage actually felt by anyone. If the work is successful we might nevertheless say that it captures rage powerfully. Perhaps it does this by depicting an enraged face. In this case it shows what rage looks like

without enabling us to perceive anyone's rage, just as a painting of a tree in my yard enables us to know how that tree looks (at least if it is at all realistic) without enabling us to perceive that tree.

On the other hand the painting might show rage by containing brush strokes looking as if they were made in the grip of rage. It is natural to imagine that these lines were made by an angry person. That is not mandatory, however; we could just as well see them as characteristic symptoms of anger and understand their angry character on these grounds, just as we could see a characteristic symptom of jaundice in a baby's yellow cornea. Angry marks on the wall look as if they were made in anger. We call the hurricane angry because it behaves as if it were someone, albeit a very large and powerful someone, in a rage. The marks and hurricane are angry, and show how anger behaves, without being tied to any particular sensory modality. Angry behavior need not have a characteristic color or sound, but rather can be characterized in such terms as that it tends to cause damage. (An object might also enable us to perceive a particular case of anger, perhaps as a result of being a photo of an angry woman. However, this would not be what makes the object expressive of anger; rather, it is the fact that it shows us how anger looks.)

Thirdly, the painting might show the anger by enabling us to know how anger feels. Perhaps it provides visual experiences congruent to anger in the sense of the foregoing sections of this chapter; or perhaps it depicts a situation which is palpably such that were we in it, we would feel anger. (We discuss this phenomenon of expression via representation in Section 7.6 below.)

The EST, then, is a cognitive theory. According to it, if a non-sentient object is angry on account of showing what anger looks like (a sculpture of an enraged man, for instance, or a livid mask) then it shows anger by enabling us to gain knowledge of how anger *looks*—likewise for other sensory modalities. If the object displays a characteristic component of anger, then it enables us to gain knowledge by showing how anger behaves. Finally, if the object shows us how anger feels, then it enables us to gain knowledge of how that emotion feels.

Crystallizing the EST, then, we may say that an *object O possesses affective or experiential quality E just in case O is a potential source of knowledge of E—either by showing how E characteristically appears, how E characteristically behaves, or how E characteristically feels*. While the EST emphasizes the ability

of an expressive object to convey knowledge, it should also be clear that the knowledge in question is at least as much a matter of knowledge how as it is a matter of knowledge that. Again, the ability of an artifact to show someone how, for instance, an emotion feels, is compatible with that person's already knowing how it feels. Further, an object might possess an affective quality, such as sadness, even though it is misleading to describe it as sad. Hence a sonata might have a sad movement, and as a result possesses sadness according to the EST account. It might still be misleading to describe the sonata as sad if it exemplifies many other emotions as well. For this reason it will be clearer to describe an object as possessing affective quality E rather than just describing it as being E.

To elucidate further what is distinctive of the EST, I'll discuss some well-known alternatives and explain why it differs from them in such a way as to be superior to them. I'll argue that the expressiveness–as–showing theory improves upon these alternatives by incorporating their insights while going beyond them, either by being more general, or by eschewing features of these theories that are not necessary. It will not be to our purpose to discuss all theories falling under either the cognitive or non-cognitive rubric. Instead I'll focus just on those that seem to me the most plausible, and discuss them to the extent that doing so enables me to bring into full relief the view I offer here.

7.5.1. Resemblance theories

As the name suggests, this view explains the affective character of an artifact like a piece of music in terms of its resemblance to human expression. The sagging and cascading sounds of a sonata resemble the behavior of a person in the throes of anguish as she weeps, moves slowly, and then throws herself to the ground. The searing guitar riff resembles a raging scream. Other music will jump, spring, and bounce just as a joyous person might do. And so on. Before considering specific versions of this theory, however, we should attend to Jerrold Levinson's (2006) argument that all theories in this category are inadequate. He offers an objection to all views of musical expression (and perhaps artistic expression generally) that depend upon a perceived resemblance between the work and an agent who is literally expressing her emotion. Levinson holds that seeing or otherwise perceiving a resemblance between A and B is never on its own a sufficient condition for seeing (hearing, etc.) A *as* B. He gives the example

of seeing the resemblance of a leafy tree and a bushy head. In seeing such a resemblance, Levinson contends, we do not *thereby* see the tree as a bushy head. Perceiving a resemblance is thus not a sufficient condition for seeing-as.

Levinson is right to point out that perceiving a resemblance is not a sufficient condition for seeing- (or otherwise perceiving-) as. The point does not, however, undermine resemblance-based views of musical (or other forms of non-sentient) expression, for it is not clear that in order to perceive the expressiveness of an object one must see it as anything other than what it is. I see the Newfoundland's face, and I see the sadness in that face, for I perceive that the face has a sad look. In order to do this, must I also see it *as* sad? That evidently depends upon how we construe the truth conditions of this locution. Does seeing α as Φ require that I imagine α to be Φ, or to be the vehicle of some agent's expression of Φ? In that case the seeing-as requirement is too strong: Surely I can perceive the sad look in the Newfoundland's face, or the contemptuous look in the face of a man whose face has been disfigured by an accident, without imagining anything at all? On the other hand if the seeing-as condition does not require use of the imagination, then for all Levinson has said, perception of a resemblance will be enough to satisfy it.

Levinson's objection to resemblance-based views does not, it seems, hit its target. Let us now consider two versions of that theory to see how they account for the relevant phenomena.

The Contour/Convention Theory Formulating what is perhaps the best known theory of musical expressiveness, Kivy's ''contour theory'', holds that ''[music's] sonic 'shape', bears a structural analogy to the heard and seen manifestations of human emotional expression'' (2002, p. 40). In listening to such music we sense its emotional characteristics by virtue of discerning this analogy. Other music, according to Kivy, has the expressive character it does as a result of conventions. The theme to *2001: A Space Odyssey* expresses ambition, and if the contour part of the theory does not account for this fact, on Kivy's view the convention part of that theory will take up the slack.

Kivy believes the discernment of resemblance between the structural features of the music and the structural features of human emotional expression is not a conscious process. For this reason he can account for

the fact that our discernment of the expressiveness of an object can be instantaneous rather than the result of a conscious calculation. In addition, Kivy is aware of the fact that a piece of music can bear a structural analogy to many other things besides expressive behavior (exploding geysers, cascading waterfalls, stampeding bison) and while on occasion program music might represent such things, music does not express them. Kivy proposes an evolutionary hypothesis to the effect that human beings are simply prone to see analogies with expressive behavior in favor of the many other analogies that they might discern, just as we are prone to see a face in an electrical plug in a wall instead of the many other things that we could see there.

According to this "contour and convention theory", then, music might have a sad sound by virtue of containing structural features isomorphic to the heard or seen structural features of a person's behavior when she is displaying her sadness. However, the sadness of a bit of music might be due either to structural features of the music, such as the development of a melody over a few measures, or because of the peculiarly melancholy sound of a chord. Kivy's "contour and convention" theory has an account to offer of the former sorts of case, but not of the latter. He is clear about this, writing,

we have yet to work one further element into the contour theory: that is the expressive chords, major, minor and diminished. These chords are generally perceived as cheerful, melancholy, and anguished, respectively … The problem is that these individual chords, not having a contour, being experienced as simple qualities, do not seem to bear any analogy at all to human behavior—hence must be expressive of cheerfulness, melancholy and anguish in some other way than that allowed by the contour theory of musical expressiveness. (2002, p. 43)

Kivy goes on to observe that there is no generally accepted explanation for the expressive features of these chords, and so the contour theory is no worse off than any other theory in this regard. Kivy infers that the fact that the contour theory cannot provide an explanation here is, as he says, "no great deficit". This is of course a fallacy. If no theory in a class C can account for a datum that is in the domain of C-type theories, it is equally possible that all theories in that class are incorrect, not that any one of them is off the hook. Leaving this aside, however, Kivy does try his hand at an explanation of the expressive qualities of chords along the following lines. He observes that the major triad C−E−G sounds stable; for instance

a movement could easily end on it. However, it would not sound at all natural to end a movement on the diminished triad C–E flat–G flat. The same is true, if to a lesser extent, of the minor triad C–E flat–G. Kivy suggests that this may be why the major triad sounds cheerful, the minor triad melancholy, and the diminished triad anguished.

These remarks are suggestive but cannot stand on their own as a solution to the problem that Kivy has raised for himself. Kivy does not tell us why cheerfulness, melancholy, and anguish should be considered in increasing order of instability. Nor is it clear why, if that were established, it would fall under the "contour" theory: In what way is instability an aspect of the contour of anguish, and is that way anything like the way in which sad behavior tends to have a drooping character that can be mapped onto a temporal progression of sounds?

While the contour/convention theory appears unable to account for the expressive qualities of such things as chords, the EST is suited to do so. The reason is that the chords C–E–G, C–E flat–G flat, and C–E flat–G each map onto a different location within the three-space we have hypothesized earlier in this chapter; each of those three locations corresponds to different emotions, and it is thus those emotions with which each of these three chords is congruent. The EST now tells us that the three chords have the expressive qualities they do by virtue of showing how each of those emotions feel.

In addition to covering more musical cases than the contour/convention theory, the EST also is applicable to areas outside music. The aforementioned angry painting, whose anger is due to its enabling us to know how that emotion feels, does not resemble, or contain components that resemble, human expressions of anger. No one depicted in the painting is angry; they are all either smugly satisfied, or are too busy fighting for survival. Or imagine a photograph that depicts a happy scene, a puppy lazing in the grass under a tree. That photograph is a non-sentient object with an affective quality. It does not resemble a characteristic expression of happiness. Again, Rodchenko's photograph, *Pioneer Girl* (Figure 7.1), is admiring not because it bears a perceptual resemblance to literal expressions of admiration, but because it enables us to know how admiration feels.

A clarification is in order. We have not disagreed with Kivy's explanation of the cases that his theory seems to handle well. We have not

Figure 7.1: Rodchenko's *Pioneer Girl* Art © Estate of Alexander Rodchenko/RAO, Moscow/VAGA, New York

challenged the claim that some sad music bears a structural analogy to the look or sound of sad human behavior, likewise for angry music, and so on. Does this mean that the EST contains a resemblance component while adding other components as well? No it does not. While the EST does not

contradict the claim that some expressive music bears a structural analogy to expressive human behavior, it does not take this as an *explanation* of the expressiveness of such music. Rather, what accounts for the expressiveness of such music is the fact that it shows how sadness sounds, looks, feels, or otherwise manifests itself. Just as a painting's resemblance to the sitter is of interest to the extent that it shows what the sitter looked like, so too the music's resemblance to manifestations of emotion is of interest to the extent that it shows us one or more dimension of those emotions.

The Emotion-Characteristics-In-Sound Theory   Steven Davies writes,

the expressiveness of music consists in its presenting emotion characteristics in its appearance … These expressive appearances … are not occurrent emotions at all. They are emergent properties of the things to which they are attributed. (Davies 1994, p. 228)

Elaborating on this, Davies explains that such expressiveness "depends mainly on a resemblance we perceive between the dynamic character of music and human movement, gait, bearing, or carriage" (Davies 1994, p. 229). Thus on this view, emotions "are heard in music as belonging to it, just as appearances of emotion are present in the bearing, gait, or deportment of our fellow humans and other creatures" (Davies 1994, p. 239).

   In a careful discussion of Davies's position, Levinson suggests a gloss of that view as follows: P is expressive of E iff P exhibits an emotion-characteristic-in-sound associated with E, that is, exhibits a sound-appearance analogous to the human emotion-characteristic-in-appearance of E (Levinson 2006). Levinson now puts pressure on Davies's theory at the following point. Everything is analogous to everything else in some respect or other. But then, how analogous does a sound appearance have to be to an emotion-characteristic-in-appearance to be relevantly associated with it, that is, to be expressive of it? Levinson contends that the only possible answer to this question is that we happen to "animate" some aural patterns in a certain way. Levinson writes,

I think it is plain that there is no answer to this except by appeal to our disposition to *hear* that emotion—rather than another, or none at all—in the music, that is, by appeal to our disposition to aurally construe the music as an instance of personal expression, perceiving the human appearances in the

musical ones, in effect animating the sounds in a certain manner … Only if this occurs does the music have the expressiveness in question, regardless of the degree of similarity between the music's appearances and the human appearances by relation to which it ends up being expressive, or alternatively, the degree of similarity between the experiences of those appearances. (Levinson 2006)

Levinson in effect argues that the only plausible elaboration of Davies's theory that answers the question, How much similarity is required to make for expressiveness? must go the way of Levinson's own position, the persona theory, to be discussed below. Davies, however, could reply with help from Kivy's hypothesis, that we tend to find some similarities more salient than others, perhaps for reasons having to do with our species' evolutionary history. In so doing, however, Davies need not suppose that those similarities that we do discern are due to seeing the music as an instance of personal expression. While it may be that by thus taking a page from Kivy, Davies will make his own theory less distinguishable from Kivy's, he will not need to postulate personae in the understanding of expressive music.

As was our attitude toward Kivy, it is no aim of the present discussion to refute Davies's own view. In fact, we might well be suspicious of an attempt to do so since as observed by Stecker 1999, it is not luminously clear just what an emotion-characteristic-in-sound actually is. Rather, if we understand Davies's own view at least enough to find it prima facie plausible, we do well to try to elucidate it, and the EST does just that. Consider an example of an emotion-characteristic-in-sound, melancholy-in-sound, or more colloquially, a melancholy sound. The EST allows us to see what that might be, that is, it allows us to see what makes some sounds melancholy, by explaining that they are like melancholy along the dimensions of intensity, pleasantness, and dynamism. What makes other sounds happy is that they are like happiness along these same three dimensions, and so forth.

The gist of the EST is that it specifies the main dimensions along which we seek resemblance between properties of non-sentient objects and sentient objects. Not just any resemblance will do; we seek affinities along particular lines. Further, because of the particular lines of affinity that the EST postulates, we can understand what an emotion-characteristic-in-sound is. An aural phenomenon P is an emotion-characteristic-in-sound

E, where E is some affective property, just in case P and emotion E are equally intense or non-intense, equally pleasant or unpleasant, and equally dynamic or static.

Finally, many of our comments about Kivy's position carry over to that of Davies. In particular, we have shown that the EST position applies to musical phenomena, particularly synchronic dimensions of music, that Kivy's position does not handle. The same goes for Davies's position. Finally, we have argued that the EST applies to cases other than music that are not touched by Kivy's theory, and the same goes for Davies's theory.

7.5.2. The Persona Theory

Another cognitivist theory we might call the *Persona Theory*. According to this view, what makes an object A have affective quality E is that when we experience it we are apt to, or are disposed to, or are invited to, imagine A to contain one or more agents expressing their E. Thus for instance what makes the symphony anguished is that when we hear it we are apt to, or are disposed to, or are invited to imagine that we hear one or more agents expressing their anguish. (We will come back in a moment to dwell on these differences of formulation.)

The most developed defense of this position is in the work of Jerry Levinson, and I shall focus on his formulation. Levinson argues first of all, as background for his position, that hearing expressiveness in music commits us to hearing it as, or to being disposed to hear it as, an expression of emotion by what he calls a "persona".

we should not consider a piece of music to be strictly *expressive* of an emotion—rather than standing in some *other*, weaker, relation to it, such as possessing a perceptual quality associated with the emotion—unless we regard it as analogous to a being endowed with sentiments capable of announcing themselves in an external manner. In short, music expresses an emotion only to the extent that we are disposed to hear it as the expression of an emotion, albeit in a non-standard manner, by a person or person-like entity.

The premise seems to be that in hearing a piece of music as expressive of an emotion we are committed to hearing that music as analogous to, that is, as similar to, a being capable of both experiencing and expressing her emotions. The conclusion that Levinson draws from this premise is that in

hearing a piece of music as expressive of an emotion we are committed to hearing it as the expression of an emotion by a person or person-like entity.

This conclusion does not follow in the absence of further assumptions concerning what would count as being relevantly analogous. The music can be analogous to a being endowed with sentiments capable of announcing themselves in an external manner, without itself being a person or person-like entity. It can be analogous to such a being by, for instance, sounding sad. Levinson does not establish that the imputation of an expressive property to music requires postulation of a persona in that music.

In spite of this shortcoming, let us turn to Levinson's theory. Here is Levinson's formulation:

a passage of music P is expressive of an emotion E if and only if P, in context, is readily heard, by a listener experienced in the genre in question, as an expression of E. Since expressing requires an expresser, this means that in so hearing the music the listener is in effect committed to hearing an agent in the music—what we can call the music's *persona*—or to at least imagining such an agent in a backgrounded manner. But this agent or persona, it must be stressed, is almost entirely indefinite, a sort of minimal person, characterized only by the emotion we hear it to be expressing and the musical gesture through which it does so.

We can agree with important components of Levinson's account of expressiveness in music while leaving others aside. First of all, the first sentence of the passage is true. However, as we have just seen, Levinson does not succeed in establishing the truth of the second sentence of this passage. For this reason we may agree with the first sentence without being committed to the view that proper appreciation of expressive music requires postulation of a persona in the music.

Although Levinson's argument for the requirement of the postulation of personae in music fails, might the conclusion of that argument nevertheless be plausible? Intuitively, it does not seem to me plausible that the postulation of personae in music is mandatory. Further, Levinson gives no reason why the postulation of personae is specifically justified in the understanding of music rather than in other artforms, or for that matter, other objects, artifacts or not, that have expressive properties. Without such a reason, the plausibility of a mandatory personae doctrine in application to things other than music is germane to the assessment of his own position. For instance, a Neopolitan Mastiff has an undeniably sad face. Surely we can discern this

without imagining a persona (canine or otherwise) expressing its sadness in or through the dog's face? So too the forest has a gloomy look to it, but I do not see the plausibility in claiming that in discerning this we are thereby committed to seeing the forest as being a persona, or as comprising personae, expressing their sadness. If such claims would not be plausible in these cases, why should they be more plausible in the instance of music?

This is not to say that the postulation of a persona in expressive music is improper. Rather, if music is readily heard as a manifestation of emotion E, a listener is *entitled*, but not *committed*, to imagine hearing an agent expressing E. She is not appreciating the music incorrectly if she imagines such an agent, but nor is she mandated so to hear the music. Similarly, there is nothing incorrect about imagining an elephant-shaped cloud to be an elephant. After all, such imagining does not have mind-to-world direction of fit, and does not commit the person doing the imagining to the claim that the cloud is an elephant or is even particularly like an elephant. So too, however, if she refrains from so imagining the cloud, she is not failing to grasp a fact about the cloud, and she is not failing to react to the cloud in a way that is obligatory. Likewise, I submit that in responding to expressive music without imagining a persona, we are not failing to grasp a fact about the music, nor failing to react to the music in a way that is obligatory.

Taking these points into account, we might reformulate Levinson's position as follows:

a passage of music P is expressive of an emotion E if and only if P, in context, is readily heard, by a listener experienced in the genre in question, as an expression of E. In so hearing the music the listener is entitled, but not obliged, to imagine hearing an agent in the music—what we can call the music's *persona*—expressing her E.

I see nothing objectionable in this modified account. However, it does not offer an explanation of *why* some passages in music are readily heard as an expression of an emotion whereas some others are not. The EST, by virtue of its elucidation of the various forms that showing can take, offers such an account. With respect to Levinson's persona theory, then, we may agree with some of its components, disagree with others, while, furthermore, offering a theory that explains why such of its components as are correct, are indeed so.

7.5.3. *The music-sounds-the-way-emotions-feel doctrine*

Carroll Pratt is famous for espousing the doctrine of a kind of iconic resemblance as between the sound of music, or at least expressive music, and the feel of emotions. His position begins with a denial that music's expressiveness is due to its arousing, or its being disposed to arouse, emotions in the listener. Pratt also denies that the expressiveness of music is due to the listener's empathetic response, citing psychological evidence that the relevant motor mimicry thought to be required for the activation of such empathy is not to be found. (New and quite different forms of evidence in favor of such motor mimicry have been found since the time that Pratt wrote.) Instead, remarking that both music and emotions are dynamic, Pratt suggests that a bit of music has emotional character E by virtue of its sharing a dynamic structure identical with an actual occurrence of emotion E in a sentient creature. This does not by itself make Pratt a resemblance theorist, although he is normally so described. His slogan could as well be read as claiming that music shows how emotions feel, where such showing is to be construed as enabling the listener to acquire or activate a skill rather than knowledge that something is so.

Pratt and Suzanne Langer cite each other's work approvingly, and for this reason the criticisms commonly applied to Langer's work may be thought to carry over to that of Pratt. However, Pratt adopts none of the semiotic trappings of Langer's position, in particular her doctrine that music is a discursive system whose symbols cannot be translated into a natural language such as English. Davies, aware of these differences, nevertheless criticizes Pratt's position as sharing some of the obscurities of Langer's, writing,

Pratt encapsulates his theory in the famous slogan, "Music sounds the way the emotions feel." To this I am inclined to respond both with "of course" and with "What is that supposed to mean?" Pratt's view, like Langer's testifies to the phenomenal character of the experience of music's expressiveness, but in so doing it fails to answer the puzzle that leads one to ask, "How is this possible?" … the phenomenal similarities between the sound of music and the "form of feelings" seems inadequate to the explanatory task. (1994, p. 136)

I propose to leave aside the question how a philosopher could find a doctrine both platitudinous and obscure. The more important point is that where Pratt's program does seem inadequate, the EST is in a position to

carry the line of explanation further. To see this, consider a footnote of Davies:

In his various writings, Pratt mentions music as possessing the character of being agitated, calm, wistful, dramatic, seductive, restless, pompous, passionate, sombre, triumphant, erotic, exhilarating, martial, pensive, languid, yearning, stately, majestic, lugubrious, ecstatic, sprightly, and aspiring. Of Pratt's examples, I find agitation, restlessness and vacillation to be those most plausibly regarded as possessing a distinctive dynamic character. But I wonder if they are properly called emotions, and if they belong with joy and triumph. (1994, p. 135)

Davies is right that, for instance, calling the music ecstatic raises all of our questions all over again. However, consider the three cases that he does think legitimate. We now know that agitation, restlessness, and vacillation do not need to be emotions in order to make music have an expressive quality. What matters is that such characteristics hold of a piece of music together with other features in such a way as to locate it in a three-space whose dimensions are the continua of pleasant/unpleasant, intense/mild, dynamic/static. Agitation and restlessness are relatively dynamic while the latter tends to be less pleasant. Vacillation tends better to characterize music that is calm for only brief periods without culminating in or recurring to any period of calm that is pleasant.

7.6. Transparency and translucency: expression in representation

Consider a photograph of a mountain scene. I shall follow Walton (1974, 1997) in holding that photos enable us quite literally to perceive what they are photographs of. Thus the photo shows us mountains by enabling us to perceive them; it is a perception-enabling form of showing. It also shows us forests atop those mountains by enabling us to perceive them. The photo, additionally, enables us to perceive a few trees, or at least their outlines, such as those atop the mountain in the middle distance. However, from the premise that I perceive a forest, and the premise that a forest comprises trees, it does not follow that I perceive each tree in the forest that I perceive. Given this photo's resolution and size, there are many trees that we know to be there but that we do not perceive. In that case the photo shows those trees by demonstrating their existence.

In addition to perception-enabling showing and demonstrative showing, the photo provides knowledge of a qualitative kind. For instance, supposing that the photo did not distort its image too much, it also shows us the color of the range. (We return in a moment to the question what to say if the camera used filters or some other manipulative device.) In this respect, too, gazing at a photo is like looking at the scene of which it is a photo. However, photographs also share their ability to show what something looks like with drawings and paintings. A painting of the same mountain range might capture its color as well as the photograph; in fact it may even do better. A drawing or painting can show how something looks, and, in general, the better it does this the more realistic it is. Observe, however, that from the premise that a painting shows how your great-grandmother looked, it does not follow that by looking at the painting you can see your great-grandmother. For this reason, in holding that a painting or drawing shows how something looks, we need not infer that it is "transparent" in Walton's sense of that term.

Representations, then, can contain qualitative and non-qualitative information, and indeed a non-photographic representation like a painting can contain qualitative information even if it does not enable viewers to perceive what it depicts. This comports with the common-sense idea that a faithful portrait painting of an ancestor can show us what she looked like. It can do so without enabling us to see her.

Representations are rich with information in other respects. We saw above that by perceiving a child getting hit by a car I may acquire knowledge of how it feels to be horrified, or that knowledge might be activated in me if it had been dormant. I can similarly acquire, or activate, knowledge of the feeling of calm as I listen to a recording of waves crashing onto a shoreline. Further, just as it is a fact of the human perceptual system that horizontal lines tend to make a thing look wider and vertical lines make it look taller, a recording of the crashing surf tends to be relaxing, and an image of something disgusting can provoke disgust.

Disgust such as this does not require the intervention of belief for its activation. I do not, in particular, need to believe the pictured object to be real in order to feel disgust as I gaze at it. Rather, that disgust can be triggered in an entirely non-doxastic way, just as the slimy slithering on my leg might trigger an automatic response of disgust, fear, or both, as I jump to rid myself of it. Some representations, then, can trigger affective

responses in us without mediation of a belief that what is represented is real or likely to be the case. In so doing they can also give us knowledge how, namely knowledge of how an experience feels.

We can also acquire knowledge how by more cognitively complex means, such as those involving counterfactual reasoning. For instance, if I can imaginatively project myself into the situation represented before me, I may also learn how I would feel in that situation were it actual. That feeling might be new to me, and I may thereby learn something from the representation of a state of affairs leaving aside the question whether it is actual. In addition, even if I am acquainted with that feeling, I might also learn from this experience that situations of *this* sort tend to provoke feelings of that kind.

Example: Imagine a painting of a squalid urban scene. No one seems to have much to do to fill up their time. Most of the storefronts are boarded up, and such establishments as there are offer ''payday'' loans at rapacious fees or liquor at midday. This painting helps one to imagine what it would be like to live in a situation like this. Not completely; those of us fortunate enough not to live like this can hardly use this photo completely to grasp the monotony, day in and day out, of living in this world. However, a sufficiently sensitive viewer may use this painting to get a sense of what such hopelessness might be like: of how it feels to have no viable prospect of economic advancement; to live amid chronic violence; to have no source of fulfillment other than drugs and drink. A series of such paintings might bring home the point with even more force. By getting a sense of these things the viewer can come to understand how *hopelessness* feels.

It is also natural to say that the painting conveys a sense of hopelessness. It has a hopeless feel to it. I suggest that it is equally true to say that the painting expresses hopelessness. By its means we do not literally perceive hopelessness. (The painting would have its expressive dimension even if we did not see any hopeless people in it.) Further, the painting does give evidence of the presence of hopelessness in or near the subject matter. However, so describing matters does not fully capture the emotional power of the painting. For just giving evidence of hopelessness does not put anyone in a position to know how hopelessness feels. Rather, the painting also *shows how* hopelessness feels, and thereby expresses hopelessness. Here is a way in which representation can have a significant expressive dimension as well.

We noted in Chapter 4 that while some forms of self-expression enable perception of what is shown, they do not enable such perception for all possible observers. Only those observers with the appropriate sensory modalities can perceive what is thereby shown. A deaf person might not hear the trepidation in my voice, while a dog might be the only creature able to perceive my anxiety. An analogous point applies to those cases of self-expression that show how something feels. Only those agents equipped to answer questions of the form, 'How would I feel were I in this situation?' will be able to gain knowledge of how something feels from the cases under discussion here. That includes creatures whose cognitive capacities prevent them from engaging in the counterfactual reasoning just sketched. Even for the cognitively sophisticated, however, one might be able to entertain that counterfactual question without having a clue what its answer might be. For a cartoon example, Mr. Spock could not use the photograph discussed above to get a sense of how hopelessness feels, since it's doubtful that he can imagine himself having emotions or feelings at all. For a more realistic example, one who for one reason or another finds it difficult to imagine herself into the world of the photograph (perhaps because of resistance; she might be distracted with such questions as, 'Why don't they just work harder?', or 'It's either eat or be eaten!') will be unlikely to learn from it how hopelessness feels.

We tend to think of ourselves being called upon to empathize with those suffering rather than those experiencing some pleasant or enjoyable emotion. That is presumably because sufferers tend to solicit our empathy more often than do others. Nevertheless it makes perfectly good sense to empathize with someone feeling a ''positive'' emotion or mood. So consider the photo from Rodchenko entitled *Pioneer Girl* (see p. 200). Notice that point of light in each of her eyes; that from her hair and the bit of her clothes you see, she's not overdressed for plowing a field or hammering railroad ties. Too, you're looking at her from below, and so it is natural to see her as large and strong. Yet the classical cut of her nose and upper lip suggest nobility. I can't but *admire* the Pioneer Girl. Correlatively, I can now empathize with the admiration that Rodchenko felt for her, or at least the admiration that the persona that his work embodies felt.

This photo shows me a lot of things, then. It shows me the girl, her scarf and tousled hair. It also shows me what *admiration feels like*. Of course

it has a polemical dimension as well, since it aims to convince me that the girl is representative of the coming workers' revolution. However, I can remain neutral on precisely what the girl represents while still admiring her combination of earthy strength and nobility.

A work of art can show me how an emotion feels, then, without that being its primary aim. Also, a work of art can show me how an emotion feels even if I am no stranger to that emotion. I might be acquainted with a certain emotion or mood although I am not able to access it consciously, through disuse, as it were. A work of art can reacquaint me with an emotion or mood by bringing it to consciousness.

Expressiveness and empathy are, then, closely linked. Something that is expressive of an emotion or experience shows how that emotion or experience appears or feels. In so doing, that thing makes know-how available to appropriately constituted and situated observers. When what has been made available is how an emotion or experience feels, such observers are then in a position to employ their imagination in such a way as to empathize with others. While expressiveness in the service of empathy is not the exclusive domain of art, and while a great deal of art aims at nothing of the kind, it nevertheless seems fair to say that *one* central function of artforms as disparate as painting, music, literature, film, and photography is that they show how emotion and experience feel in such a way as to equip us to achieve a greater rapport with others.

Appendix: Definitions and Analyses

Cue:

A *cue* is any feature of an entity that conveys information (including misinformation).

Signal:

A *signal* is any feature of an entity that conveys information (including misinformation) and that was designed for its ability to convey that information. A signal can be sent without being received, and a signal can be received without being interpreted properly, or interpreted at all.

Index:

An *index* is any signal that can only be faked with great difficulty as a result of limitations on the organism.

Handicap:

A *handicap* is a signal that can only be faked with great difficulty as a result of being very costly to produce.

Characterization of Self-Expression:

Where A is an agent and B a cognitive, affective, or experiential state of a sort to which A can have introspective access, A expresses her B if and only if A is in state B, and some action or behavior of A's both shows and signals her B.

Factual Speaker Meaning: Where P is an actual state of affairs, S factually speaker-means that P iff

1. S performs an action A intending that
2. in performing A, it be manifest that P, and that it be manifest that S intends that (2).

Objectual Speaker Meaning: S objectually speaker-means α iff

1. S performs an action A intending
2. α to be manifest, and for it to be manifest that s/he intends (2).

Illocutionary Speaker Meaning: S illocutionarily speaker-means that P φ'ly, where φ is an illocutionary force, iff

1. S performs an action A intending that
2. in performing A, it be manifest that S is committed to P under force φ, and that it be manifest that S intends that (2).

Speaker Meaning: S speaker–means something just in case S either objectually speaker–means something, factually speaker–means something, or illocutionarily speaker–means something.

The Expressiveness-As-Showing Theory: An object O possesses affective or experiential quality E just in case O is a potential source of knowledge of E—either by showing how E characteristically appears, how E characteristically behaves, or how E characteristically feels.

References

Brackets indicate the original publication date of a historically significant work.

Abrams, M.H. (1953). *The Mirror and the Lamp*. Oxford: Oxford University Press.

Adams, F., and Mele, A. (1989). 'The role of intention in intentional action'. *Canadian Journal of Philosophy* 19: 511–31.

Alberti, L.B. (1954) [1435]. *On Painting*. Trans. J.R. Spencer. New Haven: Yale University Press.

Andrew, R.J. (1963a). 'Evolution of facial expression'. *Science* 141: 1034–41.

———(1963b). 'The origin and evolution of the calls and facial expressions of the primates'. *Behaviour* 20: 1–109.

Armstrong, D. (1971). 'Meaning and communication'. *Philosophical Review* 80: 427–47.

Avramides, A. (1989). *Meaning and Mind: An Examination of the Gricean Account of Language*. Cambridge, MA: MIT Press.

Bach, K. (2001). 'You don't say?' *Synthese* 128: 15–44.

———(1994). 'Conversational impliciture'. *Mind and Language* 9: 124–62.

———and Harnish, R. (1979). *Linguistic Communication and Speech Acts*. Cambridge, MA: MIT Press.

Bar-on, D. (2004). *Speaking My Mind: Expression and Self-Knowledge*. Oxford: Oxford University Press.

Bargh, J.A., and Chartrand, T.L. (1999). 'The unbearable automaticity of being'. *American Psychologist* 54: 462–79.

Bennett, J. (1976). *Linguistic Behaviour*. Cambridge: Cambridge University Press.

Browne, J. (1985). 'Darwin and the expression of the emotions', in D. Kohn (ed.) *The Darwinian Heritage*. Princeton: Princeton University Press, pp. 307–26.

Burkhardt, R. (1985). 'Darwin on animal behaviour and evolution', in D. Kohn (ed.) *The Darwinian Heritage*. Princeton: Princeton University Press, pp. 327–65.

Carnap, R. (1996) [1935]. *Philosophy and Logical Syntax*. London: Thoemmes Press.

Carston, R. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.

Cheever, J. (1960). 'The death of Justina', in *The Stories of John Cheever* (New York: Knopf).

Chopin, K. (1972)[1899]. *The Awakening*. New York: Avon Books.

Churchland, P.M. (1979). *Scientific Realism and the Plasticity of Mind*. Cambridge University Press.

——(1989). 'Folk psychology and the explanation of human behavior,' in J. Tomberlin (ed.) *Philosophical Perspectives III: Philosophy of Mind and Action Theory*. Atascadero, CA, Ridgeview Press, pp. 225–42.

Clark, H.H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Collingwood, R. (1938). *The Principles of Art*. Oxford: Oxford University Press.

Crane, T. (1995). *The Mechanical Mind*. London: Penguin.

——(1990). 'An alleged analogy between numbers and propositions'. *Analysis* 50: 224–30.

Croce, B. (1992) [1901]. *The Aesthetic as the Science of Expression and of the Linguistic in General*, trans. C. Lyas. Cambridge: Cambridge University Press.

Darwin, C. (1877). 'The biographical sketch of an infant'. *Mind* 2: 285–94.

——1993[1958] *The Autobiography of Charles Darwin, 1809–1882, With Original Omissions Restored*, N. Barlow, (ed.). New York: W.W. Norton.

——1998[1872]. *The Expression of the Emotions in Man and Animals*, 3rd edn., P. Ekman, (ed.). Oxford: Oxford University Press.

David, P., Hingle, A., Greig, D., Rutherford, A., Pomiankowski, A., and Fowler, K. (1998). 'Male sexual ornament size but not asymmetry reflects condition in stalk-eyed flies'. *Proceedings of the Royal Society of London B* 265: 2211–16.

Davidson, D. (1989). 'Moods and performances', in his *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, pp. 109–23.

——(1974). 'Belief and the basis of meaning', in his *Inquiries*.

Davies, S. (1980). 'The expression of emotion in music'. *Mind* 89: 67–86.

——(1994). *Musical Meaning and Expression*. Ithaca, NY: Cornell University Press.

Davis, W. (2003). *Meaning, Expression and Thought*. Cambridge: Cambridge University Press.

——(1992a). 'Speaker meaning'. *Linguistics and Philosophy* 15: 223–53.

——(1992b). 'Cogitative and cognitive speaker meaning'. *Philosophical Studies* 67: 71–88.

de Jorio, A. (2000) [1832]. *Gesture in Naples and Gesture in Classical Antiquity*, trans. A. Kendon. Bloomingdon, IN: Indiana University Press.

Denkel, A. (1992). 'Natural meaning'. *Australasian Journal of Philosophy* 70: 296–306.

DePaulo, B., and Friedman, H. (1998). 'Nonverbal communication', in *Handbook of Social Psychology, Vol. II*, 4th edn., edited by D. Gilbert, S. Fiske, and G. Lindzey. New York: McGraw Hill, pp. 3–40.

——(1992). 'Nonverbal behavior and self-presentation'. *Psychological Bulletin* 111: 203–43.

Dretske, F. (1988). *Explaining Behavior: Reasons in a World of Causes*. Cambridge, MA: MIT Press.

Dummett, M. (1993). 'Mood, force and convention', in his *The Seas of Language*. Oxford: Oxford University Press, pp. 202–23.

Ekman, P. (1997). 'Expression or communication about emotion', in N. Segal, *et al* (eds.) *Uniting Psychology and Biology*. Washington, D.C.: American Psychological Association, pp. 315–38.

―――― (1993). 'Facial expression and emotion'. *American Psychologist* 48: 384–92.

―――― (1984). 'Expression and the nature of emotion', in K. Scherer and P. Ekman (eds.) *Approaches to Emotion*. Hillsdale, NJ: Erlbaum, pp. 319–43.

―――― and Friesen, W.V. (1982). 'Felt, false, and miserable smiles'. *Journal of Nonverbal Behavior* 6: 238–52.

―――― Levenson, R., and Friesen, W. (1983). 'Autonomic nervous system activity distinguishes among emotions'. *Science* 221: 1208–10.

Eliot, G. (2001). *The Mill on the Floss*. New York: Modern Library.

―――― (2000). *Middlemarch: A Study of Provincial Life*. New York: Modern Library.

Eliot, T.S. (1921). 'Hamlet and his problems', in *The Sacred Wood: Essays on Poetry and Criticism*. New York: Knopf.

Emonds, J. (1976). *A Transformational Approach to English Syntax*. New York: Academic Press.

Faigin, G. (1990). *The Artist's Complete Guide to Facial Expression*. New York: Watson-Guptil.

Field, H. (1981). Postscript to 'Mental representation', in Block (ed.) *Readings in the Philosophy of Psychology, Vol. II*. Cambridge, MA: Harvard University Press, pp. 112–4.

Frank, M., Ekman, P., and Friesen, W.V. (1993). 'Behavioral markers and recognizability of the smile of enjoyment'. *Journal of Personality and Social Psychology* 64: 83–93.

Fridlund, A. (1997). 'The new ethology of human facial expressions,' in J.A. Russell and J.M. Fernandez-Dols (eds.) *The Psychology of Facial Expression*. Cambridge: Cambridge University Press, pp. 103–29.

―――― (1996). ' "Facial expressions of emotion" and the delusion of the hermetic self', in R. Harre and W. Parrott (eds.) *The Emotions: Social, Cultural and Biological Dimensions*. Thousand, Oaks, CA: Russell Sage Foundation.

―――― (1994). *Human Facial Expression: An Evolutionary View*. New York: Academic Press.

Frijda, N. (1994). 'Emotions are functional, most of the time', in P. Eckman and R. Davidson (eds.) *The Nature of Emotion: Fundamental Questions*. Oxford: Oxford University Press.

―――― (1986) *The Emotions*. Cambridge: Cambridge University Press.

Frith, U. (1989). *Autism: Explaining the Enigma*. Oxford: Blackwell.

Gauker, C. (2001). 'Situated inference versus conversational implicature'. *Nous* 35: 163–89.

Gaut, B. (1999). 'Identification and emotion in narrative film', in Plantinga and Smith (eds.) *Passionate Views: Film, Cognition and Emotion*. Baltimore, MD: Johns Hopkins University Press, pp. 200–16.

Geach, P.T. (1976). 'Saying and showing in Frege and Wittgenstein'. *Acta Philosophica Fennica* 28: 54–70.

Gould, S.J., and Vrba, E. (1982). 'Exaptation: a missing term in the science of form.' *Paleobiology* 8: 4–15.

Green, M., (forthcoming a). 'The Handicap Principle and the expression of psychological states', forthcoming in *Mind and Language*.

—— (forthcoming b). 'Empathy, expression and what artworks have to teach', forthcoming in G. Hagberg (ed.) *Art and Ethical Inquiry*. Oxford: Blackwell.

—— (2007a). 'Expression, indication and showing what's within'. *Philosophical Studies,* 134.

—— (2007b). 'Moorean absurdity and showing what's within', in Green and Williams 2007.

—— (2003a). 'Intention and authenticity in facial expression of pain'. *Behavioral and Brain Sciences* 25: 460–1.

—— (2003b). 'Grice's frown: on meaning and expression,' in G. Meggle and C. Plunze, eds., *Saying, Meaning, Implicating*. Leipzig: University of Leipzig Press, pp. 200–19.

—— (2002). Review of Davis, *Implicature: Intention, Convention and Principle in the Failure of Gricean Theory* (Cambridge, 1998), *Philosophy and Phenomenological Research* 65: 241–4.

—— (2001). 'The inferential significance of Frege's assertion sign'. *Facta Philosophica*, 4: 201–29.

—— (2000a). 'The status of supposition'. *Nous* 34: 376–99.

—— (2000b). 'Illocutionary force and semantic content'. *Linguistics & Philosophy* 23: 435–73.

—— (1999b). 'Illocutions, implicata, and what a conversation requires'. *Pragmatics & Cognition* 7: 65–92.

—— (1999c). 'Attitude ascription's affinity to measurement'. *International Journal of Philosophical Studies* 7: 323–48.

—— (1997). 'On the autonomy of linguistic meaning'. *Mind* 106: 217–44.

—— and Williams, J. (2007a). 'Introduction', in *Moore's Paradox: New Essays on Belief, Rationality, and the First Person*. Oxford: Oxford University Press.

Grice, P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard.

—— (1982). 'Meaning revisited', in *Mutual Knowledge*, edited by N. Smith, 223–43. Academic Press. Repr. in Grice 1989.

—— (1969). 'Utterer's meaning and intentions'. *Philosophical Review* 78: 147–77. Repr. in Grice 1989.

Grice, P. (1957). 'Meaning'. *Philosophical Review* 66: 377–88. Repr. in Grice 1989.

Griffiths, P. (1997). *What Emotions Really Are*. Chicago: University of Chicago Press.

Grover, D. (1992). *A Prosentential Theory of Truth* (Princeton: Princeton University Press).

Grover, D., Camp, J., and Belnap, N. (1975). 'A prosentential theory of truth'. *Philosophical Studies* 27: 73–124.

Hager, J., and Ekman, P. (1985). 'The asymmetry of facial actions is inconsistent with models of hemispheric specialization'. *Psychophysiology* 22: 307–18.

Hare, R.M. (1989). 'Some subatomic particles of logic'. *Mind* 98: 23–37.

Harman, G. (1977). 'Review of Bennett, Linguistic Behaviour'. *Language* 53: 417–24.

——(1974). 'Review of *Meaning*, by S. Schiffer'. *Journal of Philosophy* 71: 224–9.

Harrah, D. (1994). 'On the vectoring of speech acts', in S. Tsohatzidis (ed.) *Foundations of Speech Act Theory: Philosophical and Linguistic Perspectives*. London: Routledge, pp. 374–92.

——(1980). 'On speech acts and their logic'. *Pacific Philosophical Quarterly* 61: 204–11.

Hartshorne, C. (1953). 'The monistic theory of expression'. *Journal of Philosophy* 55: 424–34.

——(1934). *The Philosophy and Psychology of Sensation*. Chicago: University of Chicago Press.

Hatfield, E., Cacioppo, J., and Rapson, R. (1994). *Emotional Contagion*. Cambridge: Cambridge University Press.

Hawthorne, N. (1990)[1850]: *The Scarlet Letter*, (ed.), with Introduction by B. Harding. Oxford: Oxford University Press.

Hess, U., and Kleck, R.E. (1994). 'The clues decoders use in attempting to differentiate emotion elicited and posed facial expressions'. *European Journal of Social Psychology* 24: 367–81.

Hinde, R. A. (1985). 'Was ''The expression of the emotions'' a misleading phrase?' *Animal Behaviour* 33: 985–92.

Hursthouse, R. (1991). 'Arational actions'. *Journal of Philosophy* 88: 57–68.

Hyslop, A. (1977). 'Grice without an audience'. *Analysis* 37: 67–9.

Isack, H., and Reyer, H. (1989). 'Honeyguides and honey-gatherers: interspecific communication in a symbiotic relationship'. *Science* 243: 1343–6.

Jeddi, E. (1970). 'Confort de contact et thermoregulation comportementale'. *Physiology and Behavior* 5: 1487–93.

Johnson-Laird, P. (1991). 'What is communication?' in *Ways of Communicating* ed. D.H. Mellor. Cambridge: Cambridge University Press, pp. 1–13.

Kearns, J. (1997). 'Propositional Logic of Supposition and Assertion'. *Notre Dame Journal of Formal Logic* 38: 325–49.

Kemmerling, A. (2002). 'Expressing an intentional state', in G. Grewendorf and G. Meggle (eds.) *Speech Acts, Mind and Social Reality*. Kluwer.

Kivy, P. (2002). *Introduction to a Philosophy of Music*. Oxford: Oxford University Press.

—— (1989). *Sound Sentiment: An Essay on the Musical Emotions*. Philadelphia, PA: Temple University Press.

Koch, P. (1983). 'Expressing emotion'. *Pacific Philosophical Quarterly* 64, 176–91.

Krebs, J., and Dawkins, R. (1984). 'Animal signals: mind reading and manipulation', in J. Krebs and N. Davies (eds.) *Behavioural Ecology*, 2nd edn. Oxford: Oxford University Press, pp. 380–402.

Langer, S. (1942). *Philosophy in a New Key*. Cambridge, MA: Harvard.

Lessing, G.E. (1949)[1766]. *Laocoön: An Essay on the Limits of Painting and Poetry*. Ed. E. A. McCormick (Indianapolis, IN: Bobbs-Merrill).

Levenson, W., Ekman, P., and Friesen, W. (1990). 'Voluntary facial action generates emotion-specific autonomic nervous system activity'. *Psychophysiology* 27: 363–84.

Levinson, J. (1990). 'Hope in the Hebrides', in his *Music, Art and Metaphysics*. Ithaca: Cornell University Press, pp. 336–75.

—— (1996). 'Musical expressiveness', in his *The Pleasures of Aesthetics*. Ithaca: Cornell University Press, pp. 90–125.

—— (2002). 'Sound, gesture, spatial imagination, and the expression of emotion in music'. *European Review of Philosophy* 5: 137–50.

—— (2006). 'Musical Expressiveness as Hearability-As-Expression', in M. Keiran (ed.) *Contemporary Debates in Aesthetics*. Oxford: Blackwell.

McDowell, J. (1980). 'Meaning, communication, and knowledge', in van Straaten (ed.) *Philosophical Subjects: Essays Presented to P.F. Strawson*. Oxford: Oxford University Press, pp. 117–39.

McNiel, D. *The Face*. London: Penguin.

Marks, L. (1978). *The Unity of the Senses: Interrelations Among the Modalities*. New York: Academic Press.

—— (1987). 'On cross-modal similarity: Auditory-visual interactions in speeded discrimination'. *Journal of Experimental Psychology: Human Perception and Performance* 13: 384–94.

—— (1982). 'Bright sneezes and dark coughs, loud sunlight and soft moonlight'. *Journal of Experimental Psychology: Human Perception and Performance* 8: 177–93.

—— (1995). 'Intermodal similarity and cross-modality matching: Coding perceptual dimensions', in R.D. Luce, M. D'Zmura, D. Hoffman, G. J. Iverson, and

A. K. Romney (eds.), *Geometric Representations of Perceptual Phenomena*. Hillsdale, NJ: Erlbaum, pp. 207–33.

———— Hammeal, R., and Bornstein, M. (1987). 'Perceiving similarity and comprehending metaphor'. *Monographs of the Society for Research in Child Development* 52.

Matravers, D. (1998) *Art and Emotion*. Oxford: Oxford University Press.

Matsumoto, D., and Lee, M. (1993). 'Consciousness, volition, and the neuropsychology of facial expressions of emotion'. *Consciousness and Cognition* 2: 237–54.

Matthews, R. (1994). 'The measure of mind'. *Mind* 103: 131–46.

Maus, F. (1988). 'Music as drama'. *Music Theory Spectrum* 10: 56–73. Reprinted in J. Robinson, (ed.), (1998) *Music and Meaning*, Ithaca: Cornell University Press, pp. 105–30.

———— (1997b). 'Narrative, drama, and emotion in instrumental music'. *Journal of Aesthetics and Art Criticism* 55: 293–303.

Maynard Smith, J., and Harper, D. (2004). *Animal Signals*. Oxford: Oxford University Press.

Mayr, E. (1974). 'Behavior programs and evolutionary strategies'. *American Scientist* 62: 650–9.

Mealey, L., Daood, C., and Krage, M., (1996). 'Enhanced memory for faces of cheaters'. *Ethology and Sociobiology* 17: 119.28.

Millikan, R. (2004). *The Varieties of Meaning*. Cambridge, MA: MIT Press.

Neale, S. (1992). 'Paul Grice and the philosophy of language'. *Linguistics and Philosophy* 15: 509–59.

Neill, A (1996). 'Empathy and (film) fiction', in Bordwell and Carroll (eds.) *Post-Theory: Reconstructing Film Studies*. Madison, WI: University of Wisconsin Press, pp. 175–94.

Ogden, C.K., and Richards, I.A. (1923). *The Meaning of Meaning*. New York: Harcourt, Brace and Co.

Packer, A. (2002). *The Dive From Clausen's Pier*. New York: Random House.

Palahniuk, C. (2001). *Choke*. New York: Anchor.

Peacocke, C. (2005). 'Joint attention: Its nature, reflexivity, and relation to common knowledge', in N. Eilan *et al* (eds.) *Joint Attention: Communication and Other Minds*. Oxford: Oxford University Press, pp. 298–324.

Peck, S. (1987). *Atlas of Facial Expression: An Account of Facial Expression for Artists, Actors, and Writers*. Oxford: Oxford University Press.

Pennebaker, J.W. (1990). *Opening Up: The Healing Power of Expressing Emotions*. New York: Guilford.

Petrie, M., Halliday, T.R., and Sanders, C. (1991). 'Peahens prefer males with elaborate trains'. *Animal Behaviour* 41: 323–31.

Pratt, C. (1931). *The Meaning of Music: A Study in Psychological Aesthetics*. New York: McGraw-Hill.

——(1952). *Music as the Language of Emotion*. Washington, D.C.: Library of Congress Lectures.

Recanati, F. (2001). 'What is said'. *Synthese* 128: 75–92.

——(1986). 'On defining communicative intentions'. *Mind and Language* 1: 213–42.

Reichert, S. (1984). 'Games spiders play III: cues underlying context-associated changes in agonistic behaviour'. *Animal Behaviour* 32: 1–15.

——(1978). 'Games spiders play: behavioral variability in territorial disputes'. *Behavioral Ecology and Sociobiology* 3: 135–62.

Reinhardt, T. (1975). 'Whose main clause? (point of view in sentences with parentheticals)'. *Harvard Studies in Syntax and Semantics* 1: 127–82.

Ridley, A. (1995). *Music, Value and the Passions*. Ithaca: Cornell University Press.

Rinn, W.E. (1984). 'The neuropsychology of facial expression'. *Psychological Bulletin* 95: 52–77.

Robinson, J. (2005). *Deeper Than Reason: Emotion and its Role in Literature, Music, and Art*. Oxford: Oxford University Press.

——(1998). 'The expression and arousal of emotion in music', in P. Alperson, (ed.), *Musical Worlds: New Directions in the Philosophy of Music*, University Park: Pennsylvania State University Press, pp. 13–22.

——(1997). 'Two concepts of expression'. *Journal of Aesthetic Education* 31: 9–17.

Rorty, R. (1965). 'Mind-body identity, privacy, and categories'. *Review of Metaphysics* 19: 24–54.

Rosefeldt, T. (2004). 'Is knowing-how simply a case of knowing-that?' *Philosophical Investigations* 27: 370–9.

Rousseau, J.-J. (1966). *Essay on the Origin of Language*, trans. A. Gode and J. Moran. Chicago: University of Chicago Press.

Schiffer, S. (1987). *Remnants of Meaning*. Cambridge, MA: MIT Press.

——(1972). *Meaning*. Oxford: Oxford University Press.

Searle, J. (1983). *Intentionality*. Cambridge: Cambridge University Press.

——(1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

——and Vanderveken, D. (1985). *Foundations of Illocutionary Logic*. Cambridge: Cambridge University Press.

Sellars, W. (1969). 'Language as thought and as communication'. *Philosophy and Phenomenological Research* 29: 506–27.

Siebel, M. (2003). 'Illocutionary force and attitude expression'. *Linguistics and Philosophy* 26: 351–66.

Smith, M. (1998). 'The possibility of philosophy of action', in J. Bransen and S. Cuypers, (eds.) *Human Action, Deliberation, and Causation*. Dordrecht: Kluwer, pp. 17–41.

Smith, W.J. (1985). 'Consistency and change in communication', in G. Zivin (ed.) *The Development of Expressive Behavior*. New York: Academic Press, pp. 51–75.

Spencer, H. (1863). *Essays: Scientific, Political, and Speculative*. London Longman.

Sperber, D., and Wilson, D. (1995). *Relevance: Communication and Cognition*, 2nd edn., Oxford: Blackwell.

Stampe, D. (1975). 'Show and tell', in B. Freed *et al* (eds.) *Forms of Representation*. Dordrecht: North-Holland: 221–45.

Stalnaker, R. (1984). *Inquiry*. Cambridge, MA: MIT Press.

Stanley, J., and Williamson, T. (2001). 'Knowing how'. *Journal of Philosophy* 98: 411–44.

Stecker, R. (2001). 'Expressiveness and expression in music and poetry'. *Journal of Aesthetics and Art Criticism* 59: 85–96.

——— (1999). 'Davies on the musical expression of emotion'. *British Journal of Aesthetics* 39: 273–81.

Strawson, P. (1964). 'Intention and convention in speech acts'. *The Philosophical Review* 73: 439–60.

Suppes, P., and Zinnes, J. (1963). 'Basic measurement theory', in D. Luce, *et al*, (eds.) *Handbook of Mathematical Psychology*. New York: Wiley, pp. 1–76.

Swoyer, C. (1987). 'The metaphysics of measurement', in J. Forge, (ed.) *Measurement, Realism, and Objectivity: Essays on Measurement in the Physical and Social Sciences*. Dordrecht: Kluwer, pp. 235–90.

Thapar, V. (1986). *Tigers: Portrait of a Predator*. London: Collins.

Tormey, A. (1971). *The Concept of Expression*. Princeton: Princeton University Press.

Tsohatzidis, S. (2004). 'Voices and noises in the theory of speech acts'. *Pragmatics and Cognition* 12: 109–55.

Turgenev, I. (1948). *Fathers and Sons*, trans. R. Hare. New York: Holt, Reinhart and Winston.

Urmson, J. (1952). 'Parenthetical verbs', reprinted in Flew (ed.) *Essays in Conceptual Analysis*. London: Macmillan.

Vanderveken, D. (1990). *Meaning and Speech Acts*, Vols. I and II. Cambridge: Cambridge University Press.

Vermazen, B. (1986). 'Expression as expression'. *Pacific Philosophical Quarterly* 67: 196–224.

Vlach, F. (1981). 'Speaker meaning'. *Linguistics and Philosophy* 4: 359–91.

Walton, K. (1997). 'On pictures and photographs: objections answered', in R. Allen and M. Smith (eds.) *Film Theory and Philosophy*. Oxford: Oxford University Press, pp. 60–75.

—— (1992). 'Seeing-in and seeing fictionally', in J. Hopkins, (ed.) *Psychoanalysis, Mind and Art*. Oxford: Blackwell, pp. 281–91.

—— (1990). *Mimesis as Make-Believe*. Cambridge, MA: Harvard University Press.

—— (1999). 'Projectivism, empathy and musical tension'. *Philosophical Topics* 26: 407–40.

Wharton, T. (2003). 'Natural pragmatics and natural codes'. *Mind and Language* 18: 447–87.

Wicker, F. (1968). 'Mapping the intersensory regions of perceptual space'. *American Journal of Psychology* 81: 178–88.

Wilkinson, G., and Reillo, P. (1994). 'Female choice response to artificial selection on an exaggerated male trait in a stalk-eyed fly'. *Proceedings of the Royal Society of London B* 255: 1–6.

Williams, A. (2003). 'Facial expression of pain: An evolutionary account'. *Behavioral and Brain Sciences* 25: 439–55.

Williamson, T. (1996). 'Knowing and asserting'. *The Philosophical Review* 105: 489–523.

Wilson, G. (1966). 'Arousal properties of red versus green'. *Perceptual and Motor Skills* 23: 947–9.

Wilson, T. (2002). *Strangers to Ourselves: Understanding the Adaptive Unconscious*. Cambridge, MA: Harvard University Press.

Wollheim, R. (1993). 'Correspondence, projective properties, and expression', in his *The Mind and Its Depths*. Harvard: Harvard University Press.

—— (1980). *Art and its Objects*, 2nd edn. Cambridge: Cambridge University Press.

—— (1968). 'Expression'. *Royal Institute of Philosophy Lectures, Vol. I*. London: Macmillan, pp. 227–44.

—— (1964). 'On expression and expressionism'. *Revue Internationale de Philosophie* 18: 270–89.

Wright, C. (1998). 'Self-knowledge: The Wittgensteinian legacy', in C. Wright, B. Smith, and C. Macdonald (eds.) *Knowing Our Own Minds*. Oxford: Oxford University Press, pp. 13–46.

Zahavi, A., and Zahavi, A. (1997). *The Handicap Principle: A Missing Piece of Darwin's Puzzle*. Oxford: Oxford University Press.

This page intentionally left blank

Index