# Chapter 1 Empathy and Emotions Unbound, Expressed and Unexpressed

**Abstract**

The emotions constitute information processing that operates in parallel with cognition (intelligence). Translation between these two differing systems occurs frequently, but emotions are not reducible to propositional (cognitive) attitudes. On background, the overall approach to the emotions of the position in this chapter is that emotions are not a natural kind: There is nothing necessarily in common between basic emotions, social pretenses, and irruptive motivational reactions ("moral sentiments"). With this background in place, the argument of this chapter is that unexpressed emotions are incomplete across all the different kinds. Empathy, as form of receptivity to the expression of emotion, implies an invitation to unexpressed emotions to attain completeness. This position is recommended to escape from the paradox that an unexpressed emotion does not exist. The occurrent but unexpressed emotion with its inchoate, emerging affective (felt) component - not a mere disposition - exists in interesting and important ways that are engaged. This position also escapes from the paradox that emotions, expressed or unexpressed, must have an affective (felt) component. Many emotions have a readily identifiable affective (felt) component, but by no means all. Three paradigm cases (and subcases) will be explored and used to drive the argument. The lack of expression is just as significant, though less obvious, than that of expression and arouses an empathic receptivity. The point is that the observation of small details, including empathic receptivity to micro expressions, informs the interpretive activity of empathic understanding providing as it were the means of a raid on the inarticulate. Additional consequences and the resulting dynamics of this discovery—unexpressed emotions are incomplete - for empathy will be explored.

## *Argument: Unexpressed Emotions are Incomplete*

The argument of this chapter is that unexpressed emotions are incomplete. The expression of an emotion is required to complete it. The short answer to why this is so is that an unexpressed emotion is vague and ambiguous without the expression that individualizes it and binds it to an occurring behavior. The long answer is delievered in the course of this chapter. This position is recommended to escape from the paradox that an unexpressed emotion does not exist. This paradox arises because the everyday

understanding of what is an emotion includes an inner feeling and an outward behavior, a biological substratum and an interhuman meaning, an irruptive aspect and a long-term dimension. Depending on where the analysis begins and the method by which it proceeds, these sometimes diverging, sometimes converging aspects of the emotions are a rich source of inconcistencies and anomalies. The occurrent emotion with its affective (felt) component - not a mere disposition - does indeed exist in interesting and important ways that will be engaged shortly. This position also escapes from the paradox that emotions, expressed or unexpressed, must have an affective (felt) component. This is a paradox for any account of the emotions that claims to apply a unitary definition. For example, irruptive anger has a strong, continuous affective dimension. In contrast, the loving relationship of a couple married for twenty years does not have one. Many emotions have a readily identifiable affective (felt) component, but by no means all.

As individuals, we do not learn to express our emotions; we learn *not* to express them; to contain them; to control them. We learn to interrupt expressing them. We learn to leave them unexpressed in the moment or actively intervene to interrupt the on-going tendency to allow the discharge along the physiological pathway laid out for the anticipated expression. This is completely consistent with smiling before a photograph is taken. Posing for the camera is distinct from expressing an emotion, though the two are sometimes combined. As will be discussed further in this chapter, both William James and Paul Ekman propose working from the outside in, putting a smile on one's face and sitting up straight as if one were in good spirits in order actually to cause a lightening and

improving of one's affective state (though there are limits to how effective this is in the face of really negative life events).

In asserting that "unexpressed emotions are incomplete" one misunderstanding should be avoided upfront. This does not mean that "unexpressed emotions are not emotions." That is not the position or argument of this chapter. However, it is a close paraphrase of the position of the logical behaviorist as represented (arguably) by Ludwig Wittgenstein (1945) or (under some interpretations) Gilbert Ryle (1949). No one is asserting that unexpressed emotions are bad in themselves. However, as will be further noted, Helen Lewis (1971), Michael Lewis (1992), and Thomas Scheff (1990) do suggest that over the long term traumatically experienced instances of shame that are unexpressed ("bypassed") can contribute to the etiology of neurosis. Simply stated, the indication is that unexpressed emotions can remain free floating and (as will be discussed in detail) can bind themselves to contingently related and unrelated expressions in mischievous, anomalous, and unpredictable ways. Make no mistake about it: Unexpressed emotions are emotions – incomplete emotions. At least three different scenarios – paradigm cases – exist in which unexpressed emotions are completed; and the major body of this chapter is designed to address the issue so I ask the reader's patience in framing matters in a stepwise, sensible way.

The emotions are and remain informational. They convey a message that discloses an evaluation of the situation that organisms, including human ones, are prepared to interpret and understand, but do so in a way that parallels and invites translation into cognitive

terms rather than actually being cognitions in the sense of requiring the deployment of concepts and experiences recognizable through categories. Philosophical narratives like those told by Ludwig Wittgenstein about how children learn the meaning of words may be properly invoked about parents and trainers substituting linguistic utterances such as "My foot hurts" or "Please" and "Thank you" for crying and stamping feet on the part of young, preverbal children expressing pain, emotion, or desire (Wittgenstein 1945). The way in which unexpressed emotions tend to "leak out" in micro expressions will have significant consequences for the operation of empathy, providing useful input to empathic processing that we might not previously have detected. This extends to a Freudian context.

The short definition of empathy is that it is a form of receptivity to the expressions of life (sensation, affect, emotion) of other individuals, especially vicarious experiences (but not limited to them), that, in turn, are interpreted with respect to the perspective of the other in the interest of understanding.[1] This is completely consistent with Max Weber's assertion: "Action that is fully experienced vicariously is empathically apparent [*Einfühlend evident*] in its experienced context of feeling" (Weber 1922: 91 (translation has been corrected against the original German)].[2] However, no sooner is empathy identified as a form of receptivity to other individual's expression of emotion than a

[1] [One published document and one unpublished document by the author are deleted to allow for anonymous review.]

[2] The German says: "Einfühlend evident ist am Handeln das in seinem erlebten *Gefühlszusammenhang* voll Nacherlebte"; and, unhappily, the translator has several words in English that do not exist in the German, e.g., "sympathetic participation." In German "nacherlebte" suggests the "after image" [*nach*] of an experience, hence, a *vicarious* experience. Happily, the German text of Weber's *Wirtschaft und Gesellschaft* (1922) is now conveniently located on the web - see: http://www.textlog.de/7306.html [URL verified on 09/07/2008].

familiar set of issues arises. Does this approach through empathy presuppose that an emotion must have a felt component? Does this approach presuppose that there is something = x that is being expressed? What about the fact, as Gilbert Ryle pointed out, that the vain man never *feels* vain?[3] Is Ryle just plain wrong? Or is he pointing to the fact that the vain man feels sad when he is ignored, happy when he is acknowledged, but that there is no specific affect, vainness that is experienced inwardly in addition to the standard emotions?

The proposal of this chapter is that an approach to these and related issues using empathy and empathic methods will enable us to make progress from a perspective that previously was insufficiently appreciated. Unexpressed emotions *live*. How and in what ways the emotion exists prior to its expression, though with certain conditions and qualifications, is the main topic. Granted, emotions also exist as dispositions—disposition are *not* the issue here. No taking the easy way out. The interest here is in occurrent emotions prior to expression, after expression, and in the process of coming to expression. The emotion is not complete until it is expressed and the unexpressed emotion is not nothing, but not quite something either. This "something that is not a nothing" is the target of inquiry.

What is this "incompleteness" of the unexpressed emotion? Consider: an unexpressed emotion is more like an open sentence or one of Frege's unsaturated concepts than like an inner private as opposed to a public object.[4] It is ambiguous and its reference vague until

[3] G. Ryle, (1949), *The Concept of Mind*. New York: Barnes & Noble, 1965: 87.
[4] For example, see G. Frege. (1892/95). "Comments on Sense and Meaning" in *Posthumous Writings: Gottlob Frege*. ed. Hans Hermes *et al*., tr. Peter Long and Roger White. Chicago: University of Chicago Press, 1979: 119.

the variable - the emerging emotional component - is bound to an individual expression such as laughing, crying, smiling, grimacing, etc. In being expressed, the emotion is bound to a particular form of expression in facial features and physiological functions and meaning about the situation such that the emotion would indeed exist in some form without these, but remain vague and ambiguous. The main point is to take the idea of an unsaturated concept as a *clue* to undercutting the thesis that an unexpressed emotion is something inherently inner. Given the comparison, the expression of the emotion is the equivalent of quantifying over the unbound variables, the emerging, inchoate affects, sensations, internal physiological milieu of pulse, heart, endocrine changes. Binding the vague and ambiguous something = x that constitutes the unexpressed emotion to specific facial, physiological, and even semantic functions, completes the emotion. "Incompleteness" does not mean there are logical truths that are not proven theorems. It means "open," "unsaturated," "unbound."

In general, other individuals have access to my emotion through the expression; whereas I have access to my emotion through its feeling. The intermediate case of my own inchoate, undifferentiated something = x is an instance where I access what I feel through its coming to expression, though what that is may not be distinct. Likewise, this case is interesting because the other accesses my feeling through empathy, though what the expression is may not be distinct. Empathy is a reminder that being expressed is a part of what the emotion is and without its expression, the emotion is indeed an emotion, but it is incomplete.

But one may object, did we not just say that the other accesses the emotion through its expression and now the emotion is supposedly unexpressed? How is that possible? Is this not a contradiction? How can we access the unexpressed through its expression? The short answer, which will be unpacked in three key cases, is that empathy homes in on what is conspicuous by its absence. It turns out that "unexpressed" is ambiguous and empathy locks onto that aspect of the emotion that is unexpressed but betrayed in the emotion. Once again, I ask the reader's patience in framing this is a stepwise fashion.

## *Background Orientation to the Emotions*

However, one may object again – this is old news. Have not Helen B. Lewis (1971), Michael Lewis (1992), Suzanne Retzinger (1987), and Thomas Scheff (1990) extensively explored unexpressed emotions, especially the paradigm case of "bypassed" shame? The short answer is "yes," and the rich distinctions offered by these authors are significant input to the process of framing the problem. Nevertheless, the basic thesis that "unexpressed emotions are incomplete" is not to be found in any of them – though arguably it is consistent with, and in some cases even implied by, statements made in these works (and the philosopher R.G. Collingwood (1938) seems to hint at it in the context of aesthetic experience). I am also asserting that the thesis brings an important conceptual tool to the conversation and that opens up research at the level of the experiential phenomenology of unexpressed emotions. As a final note, I wish to highlight that this chapter, though inspired by the work of many authors as cited, takes its orientation not from Lewis, but rather from empathy with unexpressed emotions; not from Retzinger, but from empathy with unexpressed emotions; not from Scheff or

Ekman, but rather from empathy with unexpressed emotions; not even from Collingwood, but simply from empathy with unexpressed emotions.

We have examples in which individuals hide their emotions from others, pretending everything is fine but they are not fine. The expression is suppressed and otherwise disguised. The nature of this disguise is the source of significant debate. Debate occurs over whether such dissembling is ever sufficient to fool all the people all the time. Debate occurs over the nature of what it is that remains unexpressed, what has been referred to as something = x. The experiential quality of such an emotion is highly problematic, since, under this interpretation it is not experienced but denied, repressed, suppressed, unacknowledged. For example, in the particular case of repressed shame, Helen Lewis acknowledges: "The notion of a phenomenology of bypassed shame is almost a contradiction in terms" (H. B. Lewis 1971: 233; M. Lewis 1991: 121-2).[5] This is a statement that should be generalized to any bypassed – that is, unacknowledged and therefore unexpressed – emotion. Here "bypassed" has a complex structure, but for our purposes means denied, repressed, or otherwise suppressed while still in an inchoate, emerging state so that the emotion as a something = x that is not nothing still does not have a chance to complete what it otherwise might be expected to express. Without the expression of the emotion to complete it, no felt and affective quality is left to provide the feeling aspect of the experience that is otherwise a "fast moving blur" of "strangulated affect" (Scheff 1990: 50, 88). Yet, as noted earlier, unexpressed emotions *live*; and do so in way that creates engaging difficulties for both practice and theory.

[5] Helen B. Lewis. (1971). *Shame and Guilt in Neurosis*. New York: International Universities Press, 1971. See also Michael Lewis. (1992). *Shame: The Exposed Self*. New York: The Free Press, 1995.

Yet in a related context, Freud says, “Betrayal oozes at every pore.” What a person cannot say due to individual and social constraints or conventions, shows itself in the form of slips of the tongue, parapraxes [*Fehlleistungen*] of all sorts, and, in extreme cases, psychiatric symptoms. This is precisely the point of Helen Lewis’ work on shame, which will not be repeated here, and which operates (in my opinion) without significant revisions to classic psychoanalytic metapsychology and mechanisms such as conversion, substitution, displacement, reaction formation, and condensation. Obviously Helen Lewis’ contribution includes many insights; but the area of Helen Lewis’s most powerful contribution is that shame has been neglected as a source of neurosis and social discontent. This is due in part to Freud’s turning away from his initial insight that actual seduction of the child by the parent (i.e., child abuse) was the cause of neurosis rather than the later theory of a fantasized seduction (H. B. Lewis 1987: 13, 31; also see Michael Lewis who attributes to Freud the view that the first 16 of Freud’s patients had been “sexually abused in one way or another” (M. Lewis 1990: 177)). In order to engage in the discussion that is central to my present thesis – the application of empathy to complete unexpressed emotions – it is impossible to follow up on this provocative and important position adequately here. From another perspective, “betrayal oozes at every pore” is also the point of Paul Ekman’s (2003) work on micro expressions. The unexpressed something = x also shows itself as expressions in the human face over which individuals have limited control and of which we are largely unaware. This opens whole new areas of research and is significant because it tracks occurring and occurrent

emotions that are beneath the threshold of awareness (whether due to dynamic repression or mere lack of interest is immaterial at this point).

However, before proceeding further with unexpressed emotions, the objection may be anticipated that no account of the emotions, expressed or unexpressed, has been proposed; and at least a gesture must be made in this direction. It is to that task that we now turn.

### Emotion: Not a Natural Kind

Paul Griffiths argues in *What Emotions Really Are* (1997) that no scientific evidence exists that emotions are a natural kind. Rather the evidence is to the contrary. The mere contingent accident of passivity has joined together in the ordinary word "emotion" three distinct, sometimes overlapping, occasionally incompatible sets of phenomena. Different kinds of emotions can be distinguished as (1) biologically hard-wired, reactive affective appraisal mechanisms, (2) social pretenses, and (3) irruptive affective-behavioral productions that are socially useful in motivating cooperation and keeping commitments. These affective phenomena are mutually inconsistent, overlap in some details by way of "family resemblance" (Wittgenstein 1945) but unlike families have a separate natural and social history. Do not try and boil them down to a single definition. They are not unifiable nor will they be unified by any theory. Fortunately for the thesis of this chapter, empathy is relevant to and cuts across all the different categorizations of the emotions – and indeed includes sensations such as pain as well as moods, which are not on the list. Empathy is relevant, though in different aspects and dimensions, whether the affective phenomena whooishes up from within by way of Damasio's (1999) inner physiological

milieu of the core self or is imposed by way of conforming norms and conventions from without (Hacking 1999 is *not* a constructivist but contains a statement of the latter's position that exceeds the coherence of most of the constructivist themselves).

In a paradox that will have to be clarified and of which Griffiths seem to be unaware, the first of these - Ekman's affect appraisal program (AAM) - provides information to the organism about salient events in the environment without in any obvious way being cognitions or propositional attitudes as such. Social pretenses are also problematic in that the pretense is not consciously acknowledged by the individual or the society of which he is a part, and the complete recognition of which requires an account of the unconscious. Finally, emotions are used strategically as irruptive motivations to make plausible that a person will keep his commitments. This use of the emotions requires the participation of another individual, even if only in imagination, and produces emotional reactions that are morally relevant such as shame, guilt, as well as trust and loyalty. The third instance of irruptive reactions in relation to others is where cognitions come into their own, enabling the individual imaginatively to see himself as another would.

As noted above, the evidence assembled by Darwin and further elaborated by Ekman's facial coding scheme is that many expressions of emotion are physiologically "hard-wired" into the biological infrastructure. Darwin provided a powerful description that the emotions such as astonishment, disgust, anger, suffering, high spirits, and fear have a

natural expression (1872).[6] The research of Paul Ekman (2003), setting out to prove the contrary and failing very creatively, has further confirmed that the expression of emotions is cross cultural, ("pan cultural"). [7] These results line up the work of Charles Darwin with that of Paul Ekman. Ekman has contributed significant results with what are loosely called affect appraisal programs and the automatic appraisal mechanism (AAM). Ekman noted the importance of distinguishing subtle cultural inhibitors such as display rules that, for example, inhibit the Japanese from openly contradicting perceived group norms (2003). Furthermore, Ekman designed ingenious research methods that reached across the cultural divide. He discovered that groups of New Guinea tribesmen who have never encountered a westerner, much less a TV or radio, will offer a narrative about the death of a child when shown a photograph of a person with a sad face. When shown a photo with the face of a person in fear, they will invent a story about suddenly encountering a poisonous snake. There is a small set of basic emotions such as fear, anger, sadness, happiness that are demonstrably pan cultural. Controversy exists whether certain borderline cases should be included – fear and surprise are closely related, likewise contempt and disgust – but there is substantial evidence (as assembled by Darwin and Ekman) that many of these basic emotions are shared by humans and higher mammals such as dogs and chimpanzees. We are in effect hard-wired to clench our teeth and narrow our eyes when angry, weep when suffering, open our eyes wide and pull back the edges of the mouth and shrink back bodily when afraid, smile broadly with "bright" eyes when happy. This does not mean we always do these things. But unless we catch

[6] C. Darwin. (1872). *The Expression of the Emotions in Man and Animals*. Chicago: University of Chicago Press, 1965.
[7] P. Ekman. (2003). *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. New York: Henry Holt, 2003.

ourselves in mid-act and inhibit our response - or substitute other behavior or gestures for the ordinary course of behavior - these natural dispositions are what get expressed, displayed and acted out.

A substantive issue exists over the information processing involved in the production of emotional responses. The distinction between problem solving using cognitive mechanisms versus noncognitive mechanisms, e.g., reflexes, habits, is operative. Basics (primary) affective emotions such as fear, surprise, sadness, happiness, disgust, and anger are information-processing mechanisms that operate *prior* to the formation of reportable beliefs. Even the cognitivists and hard-line propositional attitudinalists would acknowledge that marginal ("subliminal") stimuli – those beneath the threshold of conscious awareness – can produce an emotional reaction without the subject invoking a proposition in the usual, conventional sense. They must then argue that the proposition was already implicit in the reaction, or, at least, can be so translated upwards into language. For example, the triggering of fear requires recognition of a stimulus as dangerous – perceiving a poisonous snake near by – but this processing does not necessarily rise to the level of cognition. The affect is primary and the beliefs are produced concurrently (or micro seconds afterwards) in a parallel cognitive recognition of the situation and as a consequence of the primary affective response. Finding the proposition in the primitive reaction is a high bar to get over; and translating upward into language, while plausible in many contexts, opens the would-be cognition to indeterminacy of translation that impugns its accuracy as a cognition. As a result, the propositional attitude approach falls back from claiming actually to give an account of what happens with the emotions cognitively to being a mere translation of the function of

emotions into propositional (judgmental) terms instead. It is not that (basic) emotions are propositions (judgments), but rather (argues the propositional approach) that it is useful to translate them into propositions (judgments) for additional analysis. The emotional affect provides an implicit statement that can be stated in propositional terms by the translator ("theorist"), not the agent. The fall back will be like Napoleon's retreat from Moscow. This enables the analysis finally to stop trying to explain away the many counter-examples to the propositional approach of cognitive impenetrability of the emotions, ambivalence, and unconscious emotions and engage in practices that will be useful in managing the emotions such as using empathy to target unexpressed emotions in the diverse forms in which these occur.

Another group of emotions includes irruptive motivational contents and higher cognitions that contribute to the solving the commitment problem—how people of limited generosity and strong self-interest can form lasting commitments to one another and society. This is an example in which emotions are used strategically as irruptive reactions to make plausible that an individual in a community will keep the commitments formed with others.[8]

An example of keeping commitments under stress and asymmetrical information – just as in everyday life – is the classic problem of the prisoner's dilemma. Two accomplices are being interrogated separately in a classic situation of information asymmetry with a two-by-two set of options. If one of the prisoner's "confesses," defecting and implicating the

[8] R. H. Frank. (1988). *Passions within Reason: The Strategic Role of the Emotions*. New York: W. W. Norton.

other, and the other prisoner does not, then the one will be treated leniently and the hold out punished with the maximum severity. If both confess, both will be punished more than if neither confesses. If neither confesses, then they will be punished the least of all. Experience, including experimental results, indicates that in the face of the prospect of one's ex-partner confessing, it is very hard not to confess. Thus, the optimal outcome from the prisoner's point of view is hard to attain.

The point made by Frank (1988) is that the so-called higher emotions such as guilt, shame, righteous indignation (anger), regret, and trust (this list is not complete) overcome the prisoner's dilemma and restore the balance between short-term and long-term self-interest, making possible cooperation between the prisoners (and, by implication, people in competitive situations with asymmetrical information). If the prisoner's are perfectly rational, then the jailers will always be able to offer incentives that make it in the individual's self-interest (i.e., rational) to confess ("defect"). Only if the prisoners are "irrationally" loyal to one another (and their cause) can they foil the jailers. The point, of course, is this version of "irrationality" is itself a truncated and quasi-mathematical artifact that may usefully be complemented or even supplanted by an account of reasonableness that includes trust, righteous indignation, shame, etc., that enter into the calculation. In an alternative description, if the prisoners trust one another and are loyal to one another (and their cause), then they can foil the jailers. These emotions—trust and guilt—cannot simply be switched on or off. That a person does not have control over them—in that sense they are "irruptive"—means that an individual cannot easily fake

them. Hence, their usefulness in establishing and maintaining a relationship of trust – a social bond if you will – between the accomplices.

The perfectly rational jailer knows that the prisoners have emotions. It is just that the jailer is not moved by the emotions or lacks enough imagination in the ideal world scenario of the dilemma. The jailers lack the empathy to feel that anyone would find the offer of a reduced sentence unattractive in comparison with betraying the fellow prisoner's cause and his comrade. In short, the jailer lacks empathy with the prisoner and insight into why dedication to the cause and trust of one's fellow prisoner matter to the prisoners.

A slightly different angle on this is demonstrated powerfully in the Ultimatum Game (UG). In the UG one player is given a set amount of money – say twenty dollars. He is then required to hand over, at his own discretion, a portion of the money to a second player. If the second player declines the offer, then both players get zero; otherwise, they get to keep the cash according to the proposed offer. From a rational point of view, if player #1 offers two out of twenty dollars, then player #2 would still be better off taking it, since two dollars is more than zero dollars, which is what #2 had at the start. However, that is not what happens. Such offers are overwhelmingly refused. The second player forfeits his own self-interest. The question of course is Why? and What does it mean?

The standard interpretation is that the offer is interpreted as grossly unfair. This leads to "reciprocal fairness," and the sanctioning of others' unfair behavior, a sanction that also

has a negative consequence for player #2, though not as great a loss for #2 (who loses $2) and for #1 (who loses $18).

The experimenter can control for the possibility that the second player is seeking to influence subsequent episodes of the game with the same player in a self-interested way by specifying that the two people will not play with one another again. So, for example, if player #2 imagines that he will sanction the low offer from player #1 this time in order to get a more generous offer next time, the fact that these two people will not meet again should make this tactic useless.

What would we say about someone who accepted the "low ball" offer of two dollars out of twenty? Scientists have said that such people behave similarly to those who have had the portion of the brain called the "dorsolateral prefrontal cortex" (DLPFC) switched off by the application of a mild electrical current. They are rationally self-interested in a narrow definition of maximizing self-interest, quantitatively over and above how it matters to them emotionally. However, let's imagine such were not the case – the DLPFC is not switched off. What would we say? I suggest we would say that the person needed the money badly, had no leverage in the transaction, and (perhaps) suffered from low self-esteem or otherwise felt "beaten down". What was supposed to be a test of reciprocal fairness could also be used to test how accurately player #1 is able to assess the predicted response of player #2, in short, how accurate is #1's empathy based on all available input data (though I do not believe that such a test has occurred).

In order for player #1 to align on an offer that has an adequate chance of being accepted the first and only time the two players meet, above all else he has to be able to imagine how a "low ball" offer is going to make player #2 feel. The other (#2) is going to be angry at being taken advantage of. He is going to be angry that the one with the money does not have more respect for him in this situation. He is not going to feel happy or incented to cooperate. Since the offer has not yet occurred, it is hard to say that the first player knows what the second person is feeling, since the second player is feeling nothing in particular or perhaps just mild anticipation. But it is not hard to say that the first player's empathy is tested in imaging how the other individual is going to feel upon receiving the offer.

One thing is clearly demonstrated, based on experience, players in the underdog position overwhelming refuse such low offers. When asked why, they say the offers are "unfair". In a broader sense, of course, we start to get a sense that rational self-interest does extend to being treated fairly. But here fairness is left rather vague and does not align with getting two more dollars where otherwise the person gets none. What is likewise indicated is that enlightened self-interest includes empathy and is different than egoism narrowly defined.

Thus, emotions such as guilt or shame do *not* seem like transformations of simple fear or anger, on the one hand, or romantic love or jealousy, on the other. That is because they are not. The last two groupings—social pretenses and morally-relevant sentiments--include emotions that we can only experience with the participation and involvement of

other individuals. They are interrelational as such. Other than that, there is nothing—no feature or function--that these three divergent groups of emotions, distinguished by Griffiths (1997) have to have in common scientifically—not the affective content (feeling), not the behavior, or if one looks "under the hood," not the neurology or endocrinology; not the source or target; nor the object or cause. Note that these are real emotions. It is just that the category of emotion is not a natural kind and in that sense is incoherent, encompassing divergent phenomena.

Since emotions are not a natural kind but an *ad hoc* grouping, one should be able to find examples of a given emotion that fall into each of the diverging categories—affect program, social pretense, irruptive reaction. That is indeed the case. Completely varying examples of anger will display distinct aspects corresponding to each of the three distinct groups. For example, anger will be a response to driver A cutting off driver B in heavy rush hour traffic on the commute home. Anger will be righteous indignation as one reads that human rights activist Aung San Suu Kyi is still under house arrest in Myanmar. Finally, anger will show up as a simmering and slow burn that results in a calculated pay back as a person is passed over for a promotion that she feels she deserves, applies for and accepts a job elsewhere, quitting her current position on short notice. Sadness will be a basic response to a lost puppy. Sadness will be a social pretense as in mourning for a significant other in one's life with whom one has had little contact and a difficult relationship. Sadness is also a moral sentiment as when a father is disappointed – and saddened by – the behavior of his teenage son, who was caught drinking irresponsibly, watering down the vodka to replace that which he had consumed.

## Impact on Empathy of Emotions not being a Natural Kind

Without wishing to over-dramatize, when the everyday category of emotion explodes into several pieces, what happens to the function of empathy that was just described in relation to the expression of feelings (including emotions)? Does it not risk exploding too? The short answer is—yes, it does, though in an interesting way that shows us its structure and function, enabling it to serve as a form of receptivity to myriad expressions of feeling. We would expect that each would arouse, depending on our description, either completely different aspects of empathy or completely different capacities that we have unscientifically come to group together as empathy.

It is likely that only a subset of empathy will be relevant from the perspective of science for the use of the development of formal knowledge of human interrelations. So as not to create too much suspense, the components of empathy that emerge are - openness to the communicability of affect (sensation and emotion); putting oneself in the place of the other (empathically); recognizing the other as the possibility of choice and commitment and worthy of respect like oneself. These are articulated in a specific form of receptivity when the expressions of feeling are linguistically mediated, listening. Taken together, the consequences of empathy are evident. Taken together, empathy is the clearing for creating a community of shared experience. It creates community. It creates solidarity. It creates interhuman connection. It creates attachment, interpersonal bonds.

From the point of view of the basic emotions (fear, anger, sadness, happiness, and possibly surprise and disgust), the form of empathy most directly and immediately aroused is empathic receptivity. Individuals are also receptive to physical pain and forms of suffering experienced by other. All these become input to further upstream processing either in the interest of getting to know the other individual (in a strict sense of "know"), forming a social bond with the other (if one does not already exist) based on vicarious experience and empathic trace affects, or as a guide to action in relation to the other. Whether or not the basic emotion has the form of a propositional attitude, empathy targets the expression of emotion – fear, anger, sadness, happiness, etc. – as empathic receptivity *of* the other's expression. Empathy does have the form of an intentional relationship – in this case, receptivity to the other's expression – and that is so even if emotions, or some forms of emotions, do not. (Recall that the position of this chapter is that basic emotions are a parallel information processing system translatable, at least in part, into propositions, but not reducible to them. Recall also that empathy is activated by the expressions of life of real or imagined animate organisms, extending beyond emotions to include physical sensations, moods, feelings, affects as well as linguistically mediated expressions. However, the focus of this chapter is on the emotions, even in their fragmented forms, which are an important and even paradigm (set of cases) for empathy.) Furthermore, empathic receptivity reveals what the other individual is feeling as a vicarious experience. It does not necessarily disclose the cause of the emotion or its object. For example, five year old Hans refuses to go outside, and in interaction with Hans, my empathy discloses that Hans is experiencing fear. Absent additional interaction and conversation, empathic receptivity does not disclose that the cause of his fear is that

Hans expects to see a horse, and horses cause him to experience an immediate paralyzing fear. Empathic receptivity does not disclose that the object of Hans's fear is actually his father, or strictly speaking, his father's masculinity, for which the horse is a substitute through a series of associations not visible to empathy without significant further background conversation.

From the perspective of irruptive reactions, the form of empathy most directly and immediately activated is empathic understanding. This is the possibility of seeing the other as the other sees himself as seen by others. If the empathizing individual puts himself in the situation ("shoes") of the other, then it must be in such a way that the would-be empathizer quarantines her own traits and assumes those of the other (in so far as they are available to her) and then takes a further step. Given these traits, the empathizer has to consider how others then view him (the target of empathy (the empathasand)). This would extend to the possibility that the empathsand understands and "gets" (or, alternatively, does not get) that he is being empathized with by the empathizer. Note that such reciprocal empathy does not result in an infinite regress; it results in reciprocal empathy. Nor does this violate the asymmetry characteristic of empathy. Both individuals experience the satisfaction of being understood, each in her or his own way; the empathizer in being understood in her contribution to the humanness of the other and the empathasand in being understood as a fellow traveler on the road of life whose joy or suffering has been shared with the other and himself "gotten" for who he is and who he is not.

From the perspective of social pretenses, the form of empathy most directly invoked is empathic interpretation – the unmasking of the pretense as a socially and normatively sanctioned behavior that, however, is not necessarily an authentic form of relationship, self-expression, or commitment. The philosopher Gilbert Ryle famously noted that the vain man never feels vain. Ryle properly took this as a counter-example to the assertion that a feeling, agitation, or palpitation has to correspond to a given instance of emotional behavior and just as significantly that our self-understanding is incorrigible, and, even more, that self-understanding is available through others, from a third-person perspective. Naturally, the vain man feels sad when others ignore him; and the vain man feels proud when others acknowledge him. But in neither case is there an inner affect "vainness" rather than sadness or pride (etc.). This slogan – the vain man never feels vain - is able to be generalized to other arbitrary social pretenses as the rebel without a cause, twelve angry men, selected forms of romantic love (e.g., being in love with love as is Cherubino in *Le Nozze di Figaro*), the shame of the woman scorned (*The Scarlet Letter*) as well as the more obscure socially constructed emotions supposedly intrinsic to other cultures such as "wild pig" behavior (e.g., Hacking 1999).[9] Anyone who maintains that anger needs to have an object would do well to review *Rebel without a Cause*. This is a free floating position and point of view - rebellion - in search of an entire hierarchy of triggers – personal slight, cultural anomaly, social upset – to set off the interrelational running amok. In the case of the rebel without a cause, empathic interpretation discloses the social pretense as a feeling of emptiness – there is no underlying felt affect to disclose; and that is precisely what empathy discloses in this instance, the absence of feeling. The rebel

[9] Ian Hacking. (1999). *The Social Construction of What?* Cambridge, MA: Harvard University Press, 1999.

without a cause (goal) is implementing a search program to surface a cause (trigger) sufficiently significant (in this example) to cause (produce) the emotion of anger. In the case of social pretenses, distinction between the expression of the emotion and the corresponding behavior collapses. It is premature to engage in the analysis of this form since it is not even clear yet to the reader how social pretenses act as emotions. Therefore, the completion of the analysis will be deferred until the section on social pretenses as unexpressed emotions.

### The Propositional Attitude Approach to the Emotions

However, at this point it is no longer possible to delay engaging the assertion that emotions after all are propositional attitudes (sometimes also called the "cognitivist" or "judgmentalist" theory of the emotions). Fear is the proposition (e.g., belief) of the fearful individual that something is dangerous and the desire to avoid harm. "He is afraid" means "He believes he is in danger from a poisonous snake and wants to avoid the snake." If the candidate cognition that "x is a poisonous snake" is proven wrong, then the fear disappears. Taking an approach to the emotions (that they are propositional attitudes) makes giving an account of the operation of empathy relatively easy. If emotions have the structure of a propositional attitude, then the operation of empathy on the emotions is easy to explain. In expressing anger, the target of empathy, the subject, makes available the proposition to the one who is being empathic. The empathizer gets the propositional attitude, "Yes, he is angry that he was insulted." The propositional attitude is communicated multi-modally, through gesture, physiognomy, and statements (language). In expressing anger, people do not describe or report anger, they show it. The mere statement "I am angry" is unlikely to be sufficient in itself unless expressed in a

manner that suggests hostility. The recipient captures the propositional content, which, in turn, automatically evokes the associated constellation of affect, feeling, belief, desire that constitute empathizing with the other person. The "automatically" leaves open the possibility of explanatory mechanisms such as mirror neurons that lie "under the hood" in the neurological infrastructure and provide a "simulation" of the experience of the other at the level of the two organisms in resonance. Without creating too much suspense, even if emotions are not propositional attitudes, empathy will be on the critical path of understanding the expressions of others, but the ascent will be more complex (to be engaged shortly).

Unfortunately, for the proponents of the view that the basic emotions are propositional attitudes – in effect cognitions, albeit vague and obscure ones – a reduction to absurdity looms when all the details are considered. The first problem with the propositional attitude approach is the very wide net thrown over the emotions by cognition so that we can hardly escape from network, though it is we that have woven the net. Information is closely related to cognition, and many basic emotions transmit (some) information. "Snake-danger-freeze" is a nice paradigm of the informational content of fear. It is why the emotion matters to the person walking through the woods in a way that no cognition or means-end proposition ever does. The emotional reaction to the snake focuses one's attentional awareness in a way that no belief about danger or concept of survival as a goal ever could. If one looks at the definition of information, the reduction to absurdity is also available.

The fundamental definition of information is that which reduces uncertainty. According to the classic definition of information, if one is able to transmit the fundamental distinction between "on" and "off," say over a telephone line, between two systems connected by a communication channel, then you have reduced uncertainty a little bit (no pun intended).[10] This is not just a rhetorical flourish. This single distinction is sufficient to obtain information; if it is also sufficient to attain cognition, then the reduction is complete. Along similar lines, Scheffer's postulates for Boolean algebra can be derived using a single operation, a distinction indicated by the famous "Scheffer stroke." That Scheffer's operation is combination is irrelevant to the point. Though this author has not formally evaluated the claim by G. Spencer Brown (1969), the latter asserts that the propositional calculus (though not quantification theory) can be derived from a single distinction, which itself can be used to eliminate the Scheffer stroke. The point is a powerful intellectual mechanism that encompasses at least some forms of cognition can be derived from an elementary distinction as implemented in a consistent, formal notation.

Of course, the question is what is the relation between a formal definition of information and an informal definition of cognition? Let us suppose cognition is defined broadly as a distinction in experience that can be redistinguished (reidentified), then everything – even a mollusk and even an emotion - is cognitive. Since a mollusk is able to distinguish light from dark and do so repeatedly, it closes its shell when a shadow ("predator") passes over it, then you have an element of cognition. This is a reduction to absurdity of course. No

[10] C. E. Shannon. (1948). "A mathematical theory of communication," *Bell System Technical Journal,* vol. 27, pp. 379-423 and 623-656, July and October, 1948.

one imagines a mollusk has emotions, much less cognition, though it does interact with its environment in such a way as to promote its own survival. Once again a reduction to absurdity – in this case in the form of a category mistake - is pending. Emotions as propositions envision a class of propositions like no other statements with which this author is familiar.

This yields a class of propositions in which the proposition is volatile, vague, occasional, and hard to distinguish from an exclamation, exhortation, or imperative – for example, "Snake – danger – freeze" or "Snake – danger – jump away". Usually we approach propositions as being a function of synthesis and spontaneity which the speaker invokes and performs. But the propositions (judgments) as emotions are forced on us, overwhelm us. Are they propositions? Or are they ways we subsequently get access to emotions via language. The emotion is accompanied by a complex of associated propositions. The emotion is surrounded by a halo of language, and the language consists of propositions, at least in part. The weakness of this approach is the failure to allow a fundamental distinction it - a category mistake. Emotions are emotions. Propositions are propositions. Emotions are occasioned by propositions, associated with propositions, based on propositions, caused by propositions. But they are emotions, not propositions.

There is something in common sense that balks at saying we can have information without cognition (and the proposition in which the cognition is expressed). Under one interpretation, data without cognition (or a proposition) is possible—since data is just what is "given" (in Latin)—but information without cognition sounds suspicious. To say

that somethings - e.g., emotions - provide information but not cognitions gives pause. Reducing uncertainty is more than data processing; it is information processing; since we have a "yes" rather than a "no." If that is the answer to the question "Will you marry me?" then it provides quite a lot of information. Obviously the question in relation to which that represents an answer will be significant. We can have data without cognition, but if we get to information, then it is an additional accomplishment and frankly it seems like a cognitive one. So in so far as any arbitrary emotion includes a distinction – happy/sad, dangerous/trustworthy, angry/calm – we will all be cognitivists. One might say: "We are all cognitivists now!" Yet this does not seem likely to be what was intended by the original propositional attitude proposals. If we are all cognitivists, none of us are.

How can we recover the grain of truth in our sense that both emotions and cognitions are informational? A plausible first cut at cognition as distinct from information is that cognition requires predication, the application of a concept to experience and of a predicate to a subject in a statement. A table of predicates encompassing quality, quantity, relation, modality, is an informationally rich environment. Going forward that is how this chapter will distinguish those instances of information that are pre-cognitive. Instances of information that are precognitive will also be prepredicative. We may have an exclamation, exhortation, or even an imperative; but we do not have a predicate. Let us consider further consequences of the assertion that emotions are propositions.

There are further problems.

If emotions are really cognitive evaluations, then they are rather dumb ones. A less intelligent, much less intelligent, emotional mechanism has been incorporated into a more sophisticated one, the latter being primarily driven by predicate logic. Of course, the affective appraisal system, the pain sensing mechanism, and cognition are all providing information about the same environment, albeit with highly varying granularity, saliency, and partitioning of the whole. Thus, that the reactions of the emotions would agree with the judgments of cognition and the feedback from pain (sensation) receptors much of the time is perfectly intelligible and highly advantageous. However, "much of the time" is not necessarily regular or predictable. What if the emotions provide an information processing system that is parallel to but distinct from the formal system of cognition that occurs through predication in language? What if something similar can be said of the sensory system responsible for pain – that it provides an information processing system that is parallel to but distinct from both the emotions and the formal system of cognition? Then these three systems are overlapping, not mutually substitutable, or replaceable without remainder. If one could have a pre-predicative, non-propositional cognition, then it would look like an emotion. However, this changes significantly the meaning of "cognition," reducing it to a pragmatic distinction in experience (at most) and something = x (at the least). If one substitutes a broader and more capricious account of cognition as mere information processing rather than a linguistically formulable proposition, then the propositional attitude proponent (cognitivist) may be more in agreement than it at first seems.[11]

[11] Martha C. Nussbaum. (2003). *Upheavals of Thought: The Intelligence of Emotions*. Cambridge, UK: Cambridge University Press: 23.

Evolutionary psychologists (Griffiths 1997) tell us that there are trade-offs between these different information processing systems such that the emotions are quick and dirty, not that accurate, and subject to many false positives. In contrast, the cognitive system is slow and detailed in its articulation, subject to iterative improvement, and designed to put a premium on accuracy. The use of a quick and dirty system does have its advantages. The cost of not seeing a predator that is immediately present is that the individual is killed – for lunch - whereas the cost of seeing a predator that is really not there is merely a bit of wasted attention (and a fright). But notice that this is a solid account that is valid in the urban jungle, walking down a poorly lit street in a shady part of town, as well as in the early environment of evolutionary adaptation.

The empirical research of Darwin and Ekman suggests that in the case of the basic emotions we are dealing with a mechanism parallel to cognition but at a different (lower) biological level than that of human (verbal) intelligence. This results in many of the inconsistencies to which any attempt to give a single, consistent account of the emotions are subject. If the evaluation (say) of a dangerous situation is made by the cognitive mechanism pure-and-simple, then the affect will be missing. The individual will be afraid without feeling afraid, knowing that poisonous snakes are dangerous and to be avoided. If the evaluation is conveyed by the affective appraisal system, then the individual will feel afraid, but intellectually the individual may not admit that snakes are dangerous (see also Griffiths 1997: 95). As one advances beyond the basic emotions to social pretenses and irruptive moral sentiments, which do not necessarily bear the least relationship to basic emotions, then the propositional content becomes more useful in translating the emotion

into a concise, coherent statement. For example, the affect appraisal mechanism (AAM) does not accommodate all instances of anger. Many emotions are sustained responses over a long period of time, not brief episodes. As noted above, they become social pretenses ("roles")—such as rebel without a cause, twelve angry men, being in love with love ("romantic love").

### Translation between Different Information Systems – Emotions and Cognition

Basic emotions convey information. But when translated into an articulate language, the linguistic form in which the information (the distinction) is expressed looks like an imperative, exclamation or exhortation: "Snake-danger-fear-freeze!" If one were to translate and express the emotion in language as it shows up in one's sensorimotor awareness, then the result is an imperative, an exclamation, or an exhortation to oneself. "Run!" "Fight back – defend yourself!" "Jump for joy!" "Shame on you!" In some cases, it is the voice of one's parents or authority figures that is echoing in one's memory. "You should not have done that!" "Yuck!" "Guilty!" "You don't deserve it!" "You're not good enough!" The linguistic expressions into which the candidate emotions get translated do not rise to the level of predication. The linguistic forms are useful ostensively in pointing out salient features of the environment – the snake. Like the affective component, this has the result of powerfully focusing one's attention. Like pain, it is an attention grabbing mechanism with which no subject-predicate proposition, no matter how eloquent, can possibly compare in immediacy and force of what matters right here and now.

This extends to any powerful affective aspect of any emotion, though once again there is no necessary overlap between the different kinds that are loosely grouped together in ordinary language in "emotion."

Under the charitable interpretation, every emotion contains or corresponds to a proposition or cognitive appraisal. This approach points in the direction of the wisdom of the emotions. The heart knows of reasons of which the logic of the intellect is ignorant. We have not eliminated emotions in favor of propositions. Rather we have infused emotions with propositions. We have *aufgehoben* emotions into propositions. But in the course of so doing the propositional approach has lost what is essential to emotions – an immediate reaction in which the world and particular things in it is disclosed as threatening or joyous or sad or a source of anger. As in translating a poem from the language in which it was written to another natural language, nothing is lost in translation. Nothing except the poem. In translating an emotion into another language, nothing is lost in translation – nothing except the emotion. Basic emotions are and remain essentially reactive. In contrast, propositions are active, proactive, and a function of spontaneity. An analysis of "Tom is afraid that p" does not capture Tom's fear, granted that it is informative and useful in other respects. What is missing is the experience of fear.

It is useful to consider the case in which language is not available. The fussy baby – she is quite expressive but what precisely is the nature of the unhappiness – is she hot, cold, wet, teething, feverish, hungry? It is the intention of empathy that keeps the care-taker

(Mom or Dad) trying to find out even when it is by means of trial and error. Still, as an exercise, the proposal is to translate the fuss into a form of words. The baby's crying in distress translates into "I am unhappy – I want my bottle." Yet the distinction between "bottle" and "change diaper" or "close the safety pin" is contributed by the care-takers empathic listening to the quality, urgency, tone of the cry and its relation to aggregated experience with the cries of this (or similar) infants. The empathic listening of the care-taker completes the expression of the emotion. This is related, though with significant differences in detail, to what happens when the growing toddler looks at her care-taker in an act of social referencing after taking a fall in order to see whether she should laugh or cry ((Hobson 2005) more on this below).[12]

So in the single cry of the fussy infant, we have a whole network of requests, wishes, frustrated desires, satisfied longings, etc., that get expressed through the mechanism of our emotional life. The emotions start to look like a language. But not for long. When you look at the whole network of emotions, this is a language highly under-determined in its expressive precision, powers of disambiguation, and independence from the immediate context, though very powerful in its immediate, context-indexed force. We get as far as a signaling system, but not an articulated language with semantic and syntactic distinctions.

If emotions provide information about danger, delight, loss, something suddenly new, then, under this interpretation, the information conveyed can turn out to be turn or false.

[12] Peter Hobson. (2005). "What puts the jointness into joint attention?" in *Joint Attention: Communication and Other Minds: Issues in Philosophy and Psychology*, eds., N. Eilan, C. Hoerl, T. McCormack, and J. Roessler. Oxford: Clarendon Press, 2005.

We generally say that emotion is inappropriate or misfires if the poisonous snake of which I am afraid turns out to be rubber. The information provided by the emotion of fear is thoroughly pragmatic in that it enables the organism to avoid painful, possibly fatal, consequences. That the value of health and life should be dignified with the label of "truth" is altogether fitting. However, the label is transferred with the benefit of 20-20 hindsight and on the basis of parallel information processing systems, one primitive the other predicative, that conveniently converge in this case.

In making the translation of the emotion into a statement that is supposed to capture the emotion, the felt aspect of the emotion becomes fainter and indeed gets left behind as not essential. It is, however, that aspect of the emotion that was wired directly to the look on my face, palms, and mouth and the immediate reaction of freezing in my track, in the case of the snake, or preliminary palpitations in the case of the anticipated airplane take-off or gut and gag reaction in the case of worms. It is the aspect of the emotion that is qualitatively on the critical path to disclosing how and why and that the situation matters to the individual and it is why and that the emotion matters that does not get captured in the translation.

## Cognitive Impenetrability of the Emotions

If further evidence is needed that the individual is not dealing with mere belief and desire, then the evidence that emotions are impervious to beliefs is useful. This is called "cognitive impenetrability." The evidence that the individual is not really dealing with a belief that is a candidate for cognitive verification is that the individual might still feel

there is something uncanny and uncomfortable about that snake, even though it is rubber. The complex "Fear – snake – danger" does indeed contain a belief component, but it is one that becomes accessible and visible as the emotional complex gets translated into the corresponding cognitive system. For example, I *know* that flying on a commercial airplane is much safer than getting behind the wheel of a car. Yet I am still intermittently fearful, my palms get sweaty as the plane is taxiing toward the runway, and I experience a dryness in my mouth (i.e., fear) that I never experience in getting behind the wheel of a car. And I *know* which is safer - flying. Examples where true belief or knowledge does not make a difference to one's emotions are common. Darwin knew the Cobra was safely locked away back behind the plate glass in the zoo. Yet he jumped back when it struck. Of course, this is a reflex-like reaction; and it would go too far to assert that emotions are mere reflexes. They are much richer. Yet they share a reflex like function in that they are impervious to what I know (believe).

The propositional attitude theory of the emotions claims that emotions express or represent an evaluation. "Earthworms are disgusting." Yet I do not believe that earth worms are disgusting. I regard them as excellent bait when fishing down by the creek. I regard soil rich in earthworms to be rich in nutrients for growing vegetables. I consider earthworms to be quite harmless. My conscious beliefs regarding the inertness of earthworms is not taken into account in formulating my response to the sight of a mess of earthworms wiggling in the plate. My belief is overridden and I experience disgust.

The propositional attitude approach is shown to have yet another counter example by Pat Greenspan (1980). [13] Her example extends at least to the assertion that the emotions are logical in any ordinary interpretation of the principle of contradiction. The example of ambivalence where an individual holds two contrary or even contradictory emotions about a single event, outcome, individual, or situation suggests that mixed emotions do not behave as propositions in any ordinary way. Greenspan succeeds in establishing that it is really about one and the same individual that we simultaneously and with respect to the same property experience satisfaction and frustration, happiness and unhappiness, hope and fear. For this penetrating analysis, we are grateful (and without mixed feelings). Emotional ambivalence is common - love and hate, hope and fear, joy and sorrow. We have no trouble authentically entertaining mixed emotions. How this is reconcilable with the rationality of the emotions requires further argument and modifies our conception of rationality in interesting ways, according to Greenspan.[14]

## Why Emotions – and Empathy - are Indispensable

At this point yet another objection occurs. The objection is that the argument has played "fast and loose" in denying that emotions are not reducible to beliefs and desires. They are indeed so reducible. The way to lay this objection to rest is to provide a counter-example. The objector says they are so reducible; but here is an example where that is not the case.

[13] See especially P. Greenspan. (1980). "A case of mixed feelings: ambivalence and the logic of emotion" in *Explaining Emotions*, ed., A. O. Rorty. Berkeley: University of California Press, 1980: 223f.
[14] Pat Greenspan. (2002). "Practical reasoning and emotion" in A. Mele and P. Rawlings, eds., *Rationality.* New York: Oxford University Press, 2002.

Robert M. Gordon (1987) makes the case that emotions are needed as motives other than beliefs and desires (and thus not reducible to them) to explain behavior. Absent emotions, the cognitive philosopher can easily invent examples of behavior, which credibly and plausibly (though contingently) occur. These examples, though a tad artificial due to the need to constrain and control variables, are nevertheless still frequent and common enough to make the point.[15]

Gordon has considerable insight into emotions as motives. Without an account of how emotions provide us with motivations, we cannot make sense out of the behavior of people in similar situations who act differently based on (qualitatively) identical intentions and beliefs. For example, Gordon considers two farmers whose crops need rain and who have the same contingent beliefs, thoughts, and intentions, but different behavior. The one framer hopes that it will rain and the other fears that it will not rain. The first, hoping it will rain, does nothing. The other, fearing it will not rain, sets up irrigation equipment. There's no distinction in what they do without allowing for emotions.

Richard Wollheim argues against Gordon's position, but in failing to dislodge it, he actually provides further argumentation in favor of Gordon (Wollheim 1999: 107f.).[16] Decoupling the argument from how the behavior is explained—Wollheim asserts that emotions cannot function as direct motives. He then calls for the substitution of the emotion (fear) that it won't rain for the belief that it won't rain, on the one hand, and the emotion (hope) that it will rain for the belief that it will rain, on the other side. Contra

[15] Robert M. Gordon. (1987). *The Structure of Emotions: Investigations in Cognitive Philosophy*. Cambridge. UK: Cambridge University Press, 1990.
[16] R. Wollheim. (1999). *On the Emotions*, New Haven, CN: Yale University Press, 1999: 107f.

Gordon, Wollheim then attributes the view to Gordon that Gordon surreptitiously substitutes the emotion for the belief. It is simply not so. Gordon can handle maintaining different emotions alongside the same set of beliefs and desires. That is Gordon's entire point! Wollheim has to change the belief attributed to the two farmers in order to stop Gordon's argument from succeeding. Of course, in doing that, he has been unfair both to Gordon and the argument. Wollheim suggests that what the farmers really believe is not the same, contrary to Gordon. The first farmer really believes that putting out hoses will *not a make a difference* whereas the second farmer believes that it does. So, not only are the emotions different, but so are the beliefs. It is the different beliefs, not the emotions, according to Wollheim, that guide the "making a difference" of the different actions of putting out the irrigation equipment (or not) on the part of the two farmers. However, interesting this may be, it does not provide a counter example to Gordon. The "making a difference" is primarily a function of emotion, affectivity, or mood, not belief. The "difference" is not doxic (though, of course, distinctions of belief can exist, but not, by hypothesis, in this example). Wollheim just provides a different example. Gordon's analysis stands.

Even if the beliefs are different, it is common for emotions to be impervious to them (see the above discussion of the cognitive impenetrability of the emotions). Emotions are *sui generis*. Examples where true belief or knowledge does not make a difference to an individual's emotions are common. An analysis of "Tom is afraid that p" does not capture Tom's fear, granted that it is informationally useful in other ways. What is

missing is the experience of fear, the disclosure of what in the situation makes a difference – what matters – such that the situation is *threatening*, *fearsome*.

Empathy, in additional to being a fundamental condition of the possibility of human interrelations, is a practical method of accessing the expressions of emotions and other expressions of lived experience such as belief and desire as well as an important source of insight. Many of our intuitions about the subtle differences in the role of affects in different situations are empathically informed if not always empathically grounded. Our sense of when an emotion is irruptive and sudden versus sustained and enduring relies on a distinction we would not have without empathy.

If we did not have the capacity for empathy, then we would be unable to assess, gauge, or appreciate the significance of people's emotional reactions to events and to others. Such a condition exists and is sometimes called "affective blindness." Interaction with other people would loses emotional impact ("weight") and the affect blind individual loses interest in others. Such individuals are unengaged and disinterested in other people as centers of spontaneous emotional response. Of course, an individual might still value others persons if they brought groceries or washed the car. Otherwise, an individual would sit in isolation with no human contact. Without empathy we would not longer act to comfort those crying or feel a tear well up in our own eye.

Without empathy I would not be able to recognize emotions as emotions. Life would be stripped of the personal satisfaction ("meanings") that makes it worth living. While

surgical intervention at the neurological level is a rather blunt instrument, it can capture an aspect of what we mean by empathy as receptivity to the emotional expression of others. R. Joseph describes such a state resulting from surgery to the amygdala to control severe seizures. [17] The subject did well as a result of surgery, but the success was qualified. The subject reportedly no longer recognized close friends, relatives, or even his mother as targets or sources of emotional feedback. He returned indifference to his friends' anguished appeals. It is challenging to know what it is supposed to mean that this individual was otherwise cognitively unimpaired, but still did not recognize close friends. Presumably it meant that he could still do arithmetic and knew the days of the week. It would be hard to imagine a more severe cognitive impairment short of imbecility; but not because the emotions are "really" unclear cognitions; rather because of the emotional tuning that provides cognitions the basis for mattering and making a difference to the individual in the context of his or her life.

The job of education is to train the emotions to align with – make a difference in favor of - the intellect and good seeking impulses to master/control/manage the appetites. And empathy is one of the key methods in the toolkit of the philosopher, educator, and scientist, to regulate the emotions, to bring regulation of the emotions to the self. The cognitive impenetrability of the basic emotions (fear, sadness, disgust, etc.) has sometimes caused thinkers to despair about the educability of the emotions. It is a plausible next step from the assertion that the emotions cannot be educated to the alternative approach of eradicating, suppressing, neutralizing, or otherwise disabling

[17] See R. Joseph, *The Naked Neuron: Evolution of the Languages of the Brain and Body*, New York: Plenum Publishing, 1993.

them. Ignoring Socrates' sound advice as quoted by Plato that the emotions be recruited on the side of the intellect to transform the baser appetites and desires into a source of energy, the distinction between desire and emotion is collapsed and both are opposed to the intellect. Alas, the intellect is subtle, but it simply gets bowled over. In turn, this contributes to the Stoical project of mastering the passions in the interest of rationality, resulting in what Bernard Williams so eloquently calls "…the weary battle ground of reason and the emotions" (Williams 1973: 224).[18]

## Emotions are an Information Processing System Parallel to Cognition

These points telling against the propositional attitude approach would be valid even if these three systems were designed by a deity in advance and without regard to evolution or history. The third system, which delivers painful stimuli, will be set aside. Nothing depends on natural selection, variation, or evolution, which may just be a "just so story." The value of invoking evolution is as a design mechanism that accounts for the absence of Occam's razor, a principle that must be near-and-dear to the heart of every philosopher. Biological design does not follow Occam's razor and multiple, parallel information processing mechanisms have emerged where a more rational, or at least one less dependent on trail and error, would have dictated a unified, central, top down information processing mechanism. Occam's razor takes a holiday when it comes to accounting for the emotions – there are multiple, redundant systems for managing relations with the environment – the emotional one is literally quick and dirty, where

[18] Bernard Williams. (1973). *Problems of the Self*. Cambridge, UK: Cambridge University Press, 1973: 224.

"dirty" means not transparent or intellectually sophisticated, but quick – the redundancy is qualified because the cognitive (intellectual) system is much more refined, but as a result, slow.

We have hypothesized that we have at least two – possibly three - parallel information systems—the emotions as an affective appraisal system and our intellectual faculties of cognition including propositions. Can I translate every emotion into a proposition? Can I translate every emotion into a proposition without remainder? What would the remainder be or what would it look like? In the case of basic emotions what would escape from the reduction of the emotion to belief and desire would be the feeling that discloses the way the situation matters to the individual. It would also be the power of the feeling to act as a motivating force, literally causing an individual to be moved—that is, in the case of fear of the snake, to move out of the way. In the case of moral sentiments what would escape would be the pain of not living up to the individual's own standards or, just as importantly, the satisfaction of doing so. In the case of social pretenses, what would be missed is precisely the pretense. There is nothing behind the self-deceiving behavior that matters affectively. In point of fact, most emotions, occurring in the world are free of propositions (judgments) as they emerge in an individual's awareness. An engaging exercise occurs as we deploy empathy to articulate how the emotion could be translated into a form of words. Empathy has to dig beyond, behind, or between the emotions in order to find related propositions that are relevant to the context. Then empathy has to assay an articulation of what is disclosed by the emotion and how and that the situation matters to the individual.

If one thinks about it for a minute, the propositions into which one's emotions would be translated, albeit incompletely, would largely consist of profanity, four letter words, and deleted explicatives. Expressions such as "Oh heck!" and "You dope, how could you be so foolish!?" would be common. The positive emotions would be more printable, but they would not be any less excessive from a rhetorical perspective or more accurate. "You're the greatest!" Many of the propositions which really do correspond to the emotions, I submit, are experienced by individuals who do not express them and do not want to express them. They are best left unexpressed. What is more I (and most people) would want to disavow them. That is not the way I really feel about the Department Chairman. It is true that he does remind me of Rodney Dangerfield. But after all the latter was a comic genius. This leads directly to the matter of unexpressed emotions, which are pervasive in experience and life at large.

## Do we choose Our Emotions?

One important thesis of the propositional attitude approach (e.g., Solomon's (1993)) is that we are ultimately responsible for our emotions - that we have a choice about the emotions themselves distinct from the action to which a given emotion might provide a motivation. [19] If all this means is that an individual is responsible for his motives, including emotions that show up as motives, then there is no significant issue. But if this means, as it seems to imply *prima facia*, that an individual is responsible for getting back behind one's emotions and choosing them as one would choose a given course of action,

[19] R. C. Solomon. (1993). *The Passions: Emotions and the Meaning of Life*, 2nd edition. Indianapolis:Hackett.

then this is plausible and ennobling, but ultimately inaccurate. We are ultimately responsible for our behavior as based on our emotions. We are also responsible for our emotional behavior. An individual can be as angry or as miserable as she happens to be; but she is only held accountable for the actions that result from acting on those emotions. "You can be as miserable as you choose" – so we do acknowledge choice in the matter – but that choice is often simply wishful thinking. More to the point, the choice is to sit there and feel miserable, or get up and go jogging. Go to the movies. Call a friend, knowing that, in doing so, you may still feel miserable, but there is a chance you will be distracted from your misery. Doing something else creates a discontinuity in our experience. Minimally, a distraction; at its best, a new and happier experience is created.

We give ourselves feelings by imagining and rehearsing "what if" scenarios under various circumstances. This points to the circumstance that I seek out feelings, examine how I feel about something such as a proposed course of action, to help make the determination I want or wish or ought to intend to do. I also consult my feelings – that is, ask myself "What is present to my awareness? What is *there*? What am I present to in the area of feelings?" when assessing my own situation. When I experience disappointments and set backs in my goals and aspirations, then I expect to feel disappointed (sad), frustrated (angry and sad), or otherwise upset. In imagining such a set back at my job or in a relationship, I actually experience a powerful trace of the emotion itself, though one that is more fleeting and stops when I stop imagining the unwanted outcome and imagine a more favorable one. If I expect to feel bad after taking the easy way out and pocketing the waitress's tip due to poor service, then I may not do it. Instead I will leave the tip and

enjoy the satisfaction of being free of guilt over pushing down an already menial worker. Here the anticipated emotion is used as a guide to behavior. I do not choose the emotion; I choose the action based on the emotion.

We generally do not acknowledge that we *choose* our emotions. Yet many circumstances occur in which we *influence* our emotions. We rehearse a real or imagined slight in order to work ourselves up into an emotional state in which we express anger, righteous indignation or rage and demand an apology. So far these experiences are all linguistically mediated -- the expressions of emotions are asserted in statements. We go to a scary movie. What is enjoyable about being scared, knowing that we are in fact safe? The experience of the sublime – witnessing something terrifying, yet from a position of safety. We choose to go to the horror flick, which, in turn, causes the fear. We do not choose the fear as such, but rather to put ourselves in a situation known to evoke fear.[20] The experience of sudden fear represents a kind of catharsis, which is the part of the experience that we crave – a release of tension. The first *Alien* movie comes to mind with the thing you have not seen loose in the spacecraft as the ultimate "House on the Haunted Hill" scenario. Likewise, we go to a funny movie to be cheered up. The slap stick comedy, *Bringing Up Baby*, with Cary Grant and Kathryn Hepburn never fails to brighten my spirits. So some of the methods used to alter our emotions in the direction we want them to go include putting ourselves in external circumstances that incite those emotions. This is at least indirect evidence that emotions disclose specific objects in the world just as moods disclose a whole way of being in the world. However, when external

[20] Walton's assertion that this is "pretend fear" is not material to this point, which would work either way. Kendall L. Walton. (1990). *Mimesis as Make-Believe*, Cambridge, MA: Harvard University Press, 1990.

circumstances are out of line with the way we want to feel, people still have means of influencing how people feel. We tell ourselves stories. We rehearse to ourselves cheerful events or sad ones. This seems to work not just with narratives but with muscular contortions of the face. Working from the outside in, putting a smile on one's face and sitting up straight as if one were in good spirits actually causes a lightening and improving of one's affective state (though there are limits to how effective this is in the face of really negative life events). Both Paul Ekman and William James report this and recommend working from the outside inward as a way of "jump starting" – the expression is mine, not theirs – an emotional experience without, however, choosing it.

Much of our training in controlling our emotions consists in disavowing our more delicate feelings (and, in some cases, our more disruptive feelings too). As adults we really don't feel as deeply as we could and who is to say that is bad or wrong in terms of getting through the day? It is a coping mechanism to deny feelings and it works well enough when we are on "automatic pilot" going through our everyday routines by habit. When we do things habitually, we do them voluntarily but without explicit intention available to our attention. Aristotle understands that the courageous man is the man who experiences fear in the thick of threatening when it is proper to feel afraid. The courageous man rather experiences fear and goes forward into danger in spite of the fear. That is authentic courage. This is worth repeating. The courageous man does not choose the affective component of his emotion, which, in this case, is fear. The courageous man chooses to go forward into risk and danger, in spite of fear.

While an emotion resembles an intention in having an object – there is something of which I am afraid – the emotion is its own form of intentionality, not reducible to belief, desire, or perception. We lack a word or description for the kind of intentionality that occurs when an individual is the object of another's action – such as an insult- or the overwhelming force of nature – such as a tornado coming down the street on which I live. I experience myself as "at the effect" of another's intention including situations in which "intention" may be projected onto nature or into social situations that lack individuality. It is like an intention turned inside out. Heidegger had to invent a vocabulary to get back around and behind this contingency or "being at the effect of something else," calling it "thrownness" and *Befindlichkeit*.- how one finds oneself already situated in circumstances.

The above-cited debate between Gordon and Wollheim was undertaken and sustained because it was so important to understanding how emotions show up as "making a difference," as "mattering." This "making a difference" is what it means when Heidegger says the *Befindlichkeit* – his terminology for the complex of emotions, moods, sensations, affects, passions, and felt aspects of existence – is that out of which something in the world (already) matters to an individual's existence (Heidegger 1927: H137-8).[21] I am not spontaneous. I am not causally effective. Rather I am affected. I am at the effect of my affects, which disclose the world, my situation and me in it, such contingency that it matters to me and makes a difference.

[21] In der Befindlichkeit liegt existenzial eine erschliessende angewiesenheit auf Welt, aus der her Angehendes begegnen kann (H137-8). "Matters" is the *angehendes* – that which "goes towards" the individual from out of the disclosed world.

Individuals often conceal their emotions – how they really feel – because such information would give others significant insight into what the individual was going to do and how the individual was going to do it. We may be operating in circumstances of information asymmetries or even a so-called zero-sum game in which there has to be a winner and a loser. Obviously one does not want to be the loser. This opens up a distinction between the emotion and its expression. And we do speak of unexpressed emotions. We also speak of unexpressed feelings. It is not as if unexpressed sadness ceases to be sadness. I am still sad. It dos not fail to exist – say as a cause of unobservable thoughts about negative outcomes of my efforts or gloomy news reports of death or a general lethargy. But coming to expression, the unexpressed sadness undergoes a transformation. It gets expressed. For example, sometimes it becomes emotional behavior – if the person wails loudly, sobs hysterically, or rocks back and forth in visible agony. That evokes a reaction in others. How can we help? What is wrong? Shall we call an ambulance? Or it may show up as an imposition on others. What a big baby! Grow up! Get it together! Suck it up! The difference is in the narrative meaning that we recognize in the behavior.

When the feeling gets expressed, it completes the emotion. There is a sense in which an unexpressed emotion – typically a feeling or affect – is incomplete. There is something in an emotion that puts it on a trajectory towards expression in order to be whole and complete. When the feeling has been expressed, it is complete and will often disappear as an occurring phenomenon – it has been discharged. The emotion is satisfied. Often it is necessary (useful) to have another person understand what caused one to experience such

an emotion. The other is an essential part of completing the emotion, completing the expression. If no one knows what I am suffering, then I just continue to suffer it. If others know about it and share it, then I can let go of the suffering as a burden I alone must bear. Such emotions are an invitation for the empathy of others to be aroused, engaged, deployed. That such emotions sometimes evoke antipathy (as opposed to empathy), a negative reaction – "give it a break!" "cool it!" – is consistent with a privative form of empathy and even empathic distress.[22]

## *Three Key Cases and a Raid on the Inarticulate*

This argument then - that unexpressed emotions are incomplete - does not occur in a vacuum. Empathy, a form of receptivity to the expression of emotion to be further defined, implies an invitation to unexpressed emotions to attain completeness. The missing expression is just as significant, though less obvious, than that of expression and arouses an empathic receptivity. For example, that astute student of motivation and character amongst humans and related species Sherlock Holmes solves the case of the stolen race horse, Silver Blaze, in a story of the same name, by noticing that the dog did *not* bark. That is, the dog did not express itself. It might have been expected to do so upon encountering an intruder, but the thief was actually the trainer of the horse.[23] While Holmes methods are primarily inferential, he also uses empathy opportunistically: "I can

[22] This analysis occurred and was written before I had the benefit of reading Michael Lewis's *Shame* (1993), in which Lewis describes laughter, confession, and forgetting as ways of acknowledging and reducing the specific emotion of shame. With the exception of forgetting, which requires an analysis of denial, which I would analyze differently in terms of splitting off, Lewis's approach lines up with this one. The term "empathic distress" is not from Lewis – rather see Hoffman (Martin L. Hoffman. (2000). *Empathy and Moral Development: Implications for Caring and Justice.* Cambridge, UK: Cambridge University Press, 2001).

[23] See M. Truzzi, (1973), "Sherlock Holmes: Applied Social Psychologist" in *The Sign of the Three: Dupin, Holmes, Peirce*. Ed. U. Eco & T. A. Sebeok. Bloomington, IN: Indiana University Press, 1983: 64.

read in a man's eye when it is his own skin that he is frightened for."[24] The point is the observation of small details, including empathic receptivity to a micro expression of fear in the eyes, informs the interpretive activity of empathic understanding. The consequences of this discovery—unexpressed emotions are incomplete--for empathy will be explored.

However, before considering another example from literature, albeit in an essay on social psychology, a methodological point must be addressed. Literature has been used to teach the social sciences for many years, and the approach is becoming common in those areas that seek to understand agents in social interactions including various forms of dialogical in the context of human interrelations. The literary example has the methodological status of a case history, a clinical vignette, but with this one improvement. The literary author has done the phenomenological work of imaginative variations to refine the essence of the meaning functioning as an ideal type in Max Weber's sense, which, in turn, is to be further interpreted. Without wishing to rehearse the debate between understanding and explanation in the social sciences –this essay will appeal to both traditions – the use of literature here is definitely on the side of understanding.[25] For those interested, an example is provided by Thomas Scheff (1990), who uses Goethe's *Werther* (1774) in the study of shame and suicide.

[24] From "The Resident Patient," A. C. Doyle. Cited in M. Truzzi 1973: 72.
[25] Max Weber. (1922). *The Theory of Social and Economic Organization*, tr. A. M. Henderson and Talcott Parsons. New York: The Free Press, 1947. For further background see Hans H. Kögler. (2000). "Empathy, Dialogical Self, and Reflexive Interpretation: The Symbolic Source of Simulation" in *Empathy and Agency: The Problem of Understanding in the Human Sciences*, ed. Hans H. Kögler and Karsten R. Stueber. Boulder, CO: Westview Press, 2000: 191f.

As noted, three key cases are relevant. First, an unexpressed emotion where the individual is aware that he has an occurrent emotion, but makes a decision--including an implicit, default decision--not to be self-expressed in the matter. This is the simple, "control" case. Second, an unexpressed emotion where the individual is unaware that he has an occurrent emotion. This has at least three subtypes that will require analysis. The first subtype is that of an individual *denying* having the emotion (or affective component where the affect is considered prominent) in spite of the emotion having come to expression (though not display) by being betrayed in the form of a sign or symptom. For example, consider clenched teeth as a symptom of anger, short of an explosive outburst. We know the emotion is incomplete (and therefore unexpressed) because the betrayals keep occurring. The second subtype is the example of becoming aware of the *ambivalence* of the unexpressed emotion(s) in the process of expressing it (them). A third subtype is a variation on the second. An individual becomes aware that he was in a particular emotional state without it being available to his attention, but when his attention was called to it (the emotional state), then he *recognizes* and acknowledges the emotion whether or not he further expresses it. This collapses into the example of becoming aware of the emotion in the process of expressing, if he expresses it; or into key case one if he does not. The third key case is an emotion as a social pretense or role. This includes, for example, an emotion such as vanity in which case some allegedly and contingently unexpressed emotion to be surfaced is not part of the analysis since the

feeling named "vanity" does not exist, leading to the issue of what it means for vanity to be displayed, whether expressed or unexpressed.

Right up front the ambiguity of "expressing an emotion" may usefully be clarified. As used in the above three cases, "expressing" means "displaying." In general, all of the facial and physiological presentations described by Darwin in *The Expression of the Emotions in Man and Animals* and Paul Ekman in *Facial Action Coding System* are in general displays of the emotion in question whether voluntarily or involuntarily.[26] When an individual voluntarily tries *not* to display an emotion and the emotion comes out nonetheless, then the emotion is betrayed. "Unexpressed" is the complement of "expressed," covering the vast realm of non-expression. It points towards the condition necessary for a "betrayal of emotion." So unexpressed emotions are sometimes actually "betrayed" by a sign or a symptom of the unexpressed emotion such as a tic, clenched teeth, a blush. When the proper conditions are added, people speak of an unexpressed emotion being "betrayed," but people do not usually substitute "betrayed" for "displayed."

Debate exists as to whether an irruptive, reactive emotion such as anger is displayed or betrayed as it is expressed for the first time about a given situation. This author's sense of the distinction—such empathy as can be mobilized by him in the matter—is that it is a "display" even if the individual later makes the excuse that he could not help himself and just lost control. Losing control does not result in "betrayal." However, maintaining

[26] P. Ekman & W.V. Friesen, 1978), *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, CA: Consulting Psychologists Press, 2002. C. Darwin, *The Expression of the Emotions in Man and Animals* (1872), Chicago: University of Chicago Press, 1965.

insufficient control does sometimes result in betrayal. A further example will enable us to make a few preliminary distinctions.

For example, Joseph comes to ask his brothers to return home and arrives perkily and vain gloriously wearing his ceremonial burka ("coat") of many colors. With an angry roar, the brothers spontaneously jump up *en masse*, pound him with their fists, tie him up, and throw him into a nearby well.[27] The brothers are expressing their anger - anger to the point of rage - at Joseph's insufferable vanity. This expression of anger is spontaneous and unpremeditated. This is demonstrated by the subsequent discussion amongst the brothers as to what to do with Joseph now that he is in the well. They had no clue, and, running forward in imagination as to what will happen, the brothers are fearful and unhappy about the consequences when their father, Jacob, finds out.

Contrast this with what happened when, earlier in the story, Joseph perkily told his dream to his father and brothers of how the sun, the moon, and the stars bowed down to him (Joseph). Though no one's empathy rose to the level of acknowledging what happened, the brothers could audibly be heard grinding their teeth. Was this occurrence of teeth grinding an expression of emotion?

"Expressing an emotion"—in this case anger—is hopelessly ambiguous. Teeth grinding was indeed an expression of anger. All the pent up anger was bound to this one, fine-grained - one might say "micro" - response. This was evidently inadequate to reach the

[27] T. Mann. (1933/43). *Joseph and His Brothers*, tr. J.E. Woods, New York: Alfred Knopf Everyman Library, 2005: 451f.

threshold of an expression sufficiently explicit to complete the display of anger. In line with our analogy with quantification, it is just that additional open variables required binding, so the expression was still not complete, though arguably a start had been made. Rather it was a betrayal of emotion that "leaked out." The anger remained incompletely expressed and to that extent unexpressed. This had consequences later as it spontaneously erupted.

Thus, the expression of emotion includes both "acting out" and "acting from," which includes "expressing" as the high level category and does not line up neatly with "displaying" and "betraying" the emotion. In both cases, the brothers are "acting out." They spontaneously jump up and pound Joseph upon seeing him. In contrast in the earlier situation as Joseph tells his dream, the brothers push down their anger in Jacob's presence but it leaks out symptomatically—they give off a sign by grinding their teeth. A different example is required to see "acting from," which, of course, changes the story. Had the brothers performed some planning and written a letter to Jacob saying, for example: "We are sorely vexed with Joseph, but you, Dear Father, are the cause, since you are playing favorites," then the brothers would have been acting *from* anger, but not acting *out* anger (note the difference between "acting out anger" and "acting out *of* anger," the later being a synonym for "acting from anger"). In this case, writing the letter is also properly described as an aim-inhibited transformation of the anger. It makes the point, but does so in a way that follows socially defined roles such as the Ten Commandments, which, in terms of the time-line of the story, had actually not yet been delivered. The brothers are indeed truly expressing their anger, acting from anger.

However, as the story is told, the brothers wrote no letter, had no plan, and that makes the story all the more engaging. This also raises the possibility that the same behavior "teeth mashing" can in one occasion be an expression of anger - surely the brothers clenched their teeth as they pummeled the boy, though it was not the main expression of their anger - and in another a betrayal of the emotion. For example, if Ruben says, "You should not have said you dreamt that the sun, moon, and stars bowed down to you" with teeth clenched and as a habitual thrower of large rocks that are ready at hand (but without throwing any rocks), then that is an expression of emotion – a warning to be exact, that Ruben is about to blow a gasket and act out. But if Ruben throws the rock, then he is not only expressing anger, he is displaying it. He is showing it, not just saying it, not symbolically representing it, nor presenting a sign or symptom of anger. He is acting out anger, and acting out of anger. Naturally, hitting out and landing a blow, say, on Joseph's head, is not a symbolic expression of anger – it is the full blown thing itself in all its occurrent fury.

In either case, acting from or acting out, the anger does *not* cause the punch to be thrown. Rather Joseph's vanity, the brothers' being disinherited by Jacob, combined with their lack of impulse control, etc. *do* cause it. In the example of a perceived insult followed by physical retaliation, the cause of the punch is not anger; rather it is the insult and the motivation of retaliation. As an irruptive reaction, the anger occurs simultaneously with the physical affects and bodily transformations that are (identical proposition) anger. To that extent, the James-Lange theory (1884) avoids the faulty causal analysis that the emotion precedes its expression only to commit an equally egregious blunder that the

physiological changes are the cause of the emotion.[28] In this case, the physiological changes, the contingently hidden physiological milieu of heart beat, pulse, endocrine secretion, are an example of the something = x that is not nothing. With these basic distinctions handled, we now turn to the three key cases.

### Key Case One: Contingently Unexpressed Emotions

The first key case is that of unexpressed emotion where the individual is aware that he has an occurrent emotion, but makes a decision--including an implicit, default decision--not to be self-expressed in the matter.

The first key case includes the unexpressed love (affection) for a family member (parent, sibling), friend, or even competitor or opponent, who becomes unavailable through death or loss of contact over a period of years The emotion is incomplete and the unexpressed love for the mother (abandoned by the son who never looked back after a dispute) takes on the nature of what Collingwood (1938) described as a perturbation or "feeling of oppression."[29] Contra Collingwood, this is *not* a case where a person has to express his or her emotion in order to come to awareness ("knowledge") of what it is, though there are such instances as we shall see. The person is completely aware that her or his mother is an object of supreme affection (i.e., love), but never gets around to expressing it.

[28] W. James. (1884). "What is an emotion?" *Mind* 9: 188-205.

[29] R.G. Collingwood. (1938). *Principles of Art*. Oxford University, 1958: 117.

The unexpressed emotion is incomplete and lives on. It lives on as a motivating force for regret, imaginary restitution with other mother-like individuals (which might, in general, be good), flattening of affect in related areas, or "if only I had done so" conversations with friends and therapists. Some of these conversations about the unexpressed love (say) for mom can put the unexpressed emotion on the path to completion, but, they will likely have the initial character of a substitution, simulation, or second best effort, requiring work and rework.

If a simpler first key case is useful, then consider anger over a perceived or imagined slight. For example, a person is the target of a perhaps well-intentioned, but not appreciated, compliment that goes awry. The anger is stuffed down under a display rule such as a mild smile—the Japanese are reportedly good at the innocuous, uninformative mild smile (Ekman 2003). In this example, the suggestion is that the individual would acknowledge, even if grudgingly, that he was angry if asked later by a trusted friend. If not, however, then this example actually belongs in case number two (below) where the individual has an unexpressed emotion and is not aware of it. In fact, denies it.

In either case, the unexpressed emotion lives on as a motive in mis-judging and de-valuating the would-be compliment payer in a demeaning or disrespectful way. This does not have to happen; but it often does, becoming the source of cynical humor, "just get over it." In a worst case scenario, small slights can result in large breakdowns in emotional equilibrium in an individual or a group. This does not mean one should respond uninhibitedly to the expression of emotion by expressing one's own exaggerated

emotional reactions. Keeping in mind the above-cited distinctions between "acting out" and "acting from," completing the expression of one's unexpressed emotion into an empathic receptivity by means of aim-inhibited acting from the emotion allows the emotion to be completed, creating a clearing for community.

One may object that expressing one's hatred for another does not clear the way to community. Three related answers are available. First, hatred is a powerful bound between individuals and peoples. The result is a community, but not one that is functional. This is not satisfactory. Second, empathy with negative emotions, affects, feelings, and sentiments of all kinds is possible, but a relatively rare occurrence. Usually what is occurring is acting out hatred (and in effect betraying hatred), not expressing it into a rich empathic listening that allows for its completion. Thus, Joseph's brothers are acting out their deep-seated animosity ("hatred") when they beat him up and throw him into the pit. The constellation of unexpressed emotions remains incomplete, and indeed, if you read the remainder of the story, festers for years until the brothers are able to get back in communication with Joseph and complete the relationship. Thus, third, hatred when expressed into a rich empathic listening that allows for its recognition and completion does open the way to community in a positive sense. Thus, when Palestinians and Israelis (Jews), Hindus and Moslems, English Protestants and Irish Catholics, are brought together in a forum for understanding and give expression to what which they simply cannot ever, ever, ever forgive – i.e., the source of deep-seated hatred – then the possibility of community opens up. This is not to say that it is easy. This is not to say that such encounters are always guaranteed to work. This is not to say that my abbreviated

account is adequate to the matter. This is just to say that the expression of hatred (and other negative feelings) is not a counter-example, and certainly not necessarily so.

There is one intermediate case that deserves discussion prior to moving on to key case two. In early July 1863, President Abraham Lincoln learned of a great victory of the Union Army over previously unstoppable General Lee in a battle at Gettysburg, PA. Lincoln was happy about this victory. A couple of days later, he learned that Lee's army was trapped by the flood swollen Potomac River with Meade's Union Army near by. Lincoln was not just happy. He was as near to ecstatic as pensive Ole Abe ever got. Then the next day Lincoln experienced different emotions. Meade had not taken action. Lee escaped. Lincoln sat down and wrote one of his most famous documents—after the Gettysburg address and his letter of condolences to Mrs. Bixby who had lost five sons in the conflict. In it he stated his distress and disappointment to General Meade in direct and professional terms. Lincoln never sent the letter, and it was discovered in his desk after his death.[30] Were his emotions expressed? Had he merely thought about it in bed before falling asleep, one could justify saying "No." But actually having taken pen to paper, the expression was complete. There was a slight delay—and other details to which we do not have visibility—but Meade soon had a new boss, General Grant. Self-expression does indeed count for completion, but the self-expression must be robust, re-identifiable, and merely contingently unavailable to others. Strictly speaking, the person has discovered how he feels in a process of self-expression that, due to contingent circumstances, excludes expression for the benefit of another.

[30] Cited in Dale Carnegie. (1936). *How to Win Friends and Influence People*. New York: Simon & Schuster, 1981: 39-40.

Another borderline case occurs as a person expresses his feelings without thereby recognizing what they are. Here it is typical for the person to deny that he is feeling thus-and-so when one mentions the name of the feeling to him. This contains the nucleus of an account of dynamically unconscious feelings where, due to resistance, a person is unable to recognize a feeling for what it is. Frequently what the person does is in marked contrast to what he avows.[31]

In section 3 on unconscious emotions of Freud's 1915 paper "The Unconscious," he explicitly says that after repression an "unconscious affect" is only present as "a potential disposition" in the "system unconscious" whereas an unconscious idea has an actual, not dispositional, representation there. In contrast, in the section on "affect" in *The Interpretation of Dreams*, Freud notes that if an individual dreams of being afraid, then the person is really afraid. [32] The robbers are imaginary, the fear is real. If Hans dreams about a horse, the horse may really be a substitute for his father (or his father's powerful masculinity). However, if Hans, the person, is afraid, he is really afraid. The fear is not a substitute for some other emotion. It is fear. But *of what* is the issue. Thus, when an emotion shows up in a dream, day dream, or ordinary awareness, it is exactly what it seems to be as an affect (feeling). This is why empathically targeting ("locking on") the emotion—especially if the emotion is ambivalent or unexpressed—is so powerful. The emotion is occurrent but unbounded as to its cause or target and can provide clues as to

[31] S. Freud. (1915). "The unconscious" in *General Psychological Theory*, ed. P. Rieff. New York: Collier, 1963: 116-50.
[32] S. Freud. (1900). *The Interpretation of Dreams,* tr. J. Strachery. New York: Avon Books, 1965: 497f.

the individual's motivations amidst otherwise disguised and distorted expressions and behavior.

In fact, the possibility of the expression of emotion of which a person is unaware and inclined to deny it when confronted with it ("dynamically unconscious") is one reason why an emotion should not be reduced to the mere inner feeling. This represents an example of the strongest sense in which an expression of unexpressed emotion occurs without being felt as an affect available for attentional monitoring. This leads immediately to the next case.

### Key Case Two: Subtype Denial

The second key case focuses on an unexpressed emotion where the individual is unaware that he has an occurrent emotion. This has at least three subtypes that will require analysis. The first subtype is denying having the emotion (or affective component where the affect is considered essential) in spite of the emotion having come to expression (though not display) by being betrayed in the form of a sign or symptom. We suspect the emotion is incomplete (and therefore unexpressed) because the betrayals keep occurring.

For example, a couple in marital therapy puts on happy faces towards one another, the therapist, and the recording video camera. But a micro expression of contempt flickers across his face for a fraction of a second. A micro expression is a superset of the betrayal of a feeling or emotion. This is visible when the video is slowed down; otherwise it is beneath the threshold of conscious perception. The couple acknowledges having issues

(problems), but asserts respect for one another, and denies contempt. The latter is present, according to the video, and is predictive, within a rigorous margin of error, that the marriage is in trouble and will end in divorce with a probability better than chance.[33]

The above-cited discussion about by-passed shame (and unacknowledged emotions in general) as well as Freud's position in the matter should be recalled at this point. For those subjected to what Helen Lewis (1971) describe as a shaming environment, where the traumatic emotion has occurred repeatedly over an extended period of time, the unexpressed shame has become a persisting disposition and character trait. In the initial shaming environment – say the first three times it happened – the individual's actual lack of control over the situation – whether due to internal inhibition, external force, or both – was such as to cause what, for all we know (and since we will never have a video), could accurately be described as emerging, inchoate shame to be bypassed and remain something = x. Earlier in this essay, Griffiths's approach to the emotions was laid out, indicating that obvious emotions like anger and less obvious emotions such as social pretenses (vanity) belonged to different natural kinds, and were only accidentally able to be combined under the vernacular label "emotion." That position stands fact; but it may usefully be refined in an important detail. After an occurrent emotion has been repeated sufficiently often, the experience of it is imprinted as a character trait, perhaps required to stay alive at all, and comes to resemble a social role relevant to one individual and for his or her life. This is indeed a contingent transformation between the individual occurrences

[33] S. Carere and J. Gottman. (1999). "Predicting divorce among newlyweds from the first three minutes of a marital conflict discussion," *Family Process* 38, no. 3 (1999): 293-301 cited in M. Gladwell. (2005). *Blink*. New York: Little & Brown 2005: 246. See P. Ekman, (1985), *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage*. New York: W.W. Norton, 1985.

of shaming and the social pretense of the "shamed individual as a way of life." One example of the latter is Hester Prynne and the scarlet letter, in the Nathaniel Hawthorne novel of the same name. In most of the examples with which Helen Lewis is concerned, the individual wears the scarlet letter in their heart, and that presents a different dynamic from Hawthorne's.

Fast forward ten or twenty years. Is the emotion – in this case, shame - occurrent? Yes and no. It is occurrent if and when therapy (or any relevant conversation in which a generous and gracious empathic listening is present) succeeds in evoking the quality of the original experience in order to work through the original trauma, express what has long remained unexpressed, and complete it. Across such a long distance and so many bypasses, this raid on the inarticulate may or may not succeed. The importance of evoking the situation of the original trauamatization in a supportive milieu of empathic receptivity and understanding is on the critical path to completing the unexpressed emotion.

Note well that no one is asserting that an unexpressed, inner feeling necessarily must betray itself—not Darwin, Ekman, or this author. Experience shows that unexpressed emotions tend to come out. However, this is a contingent, useful statistical regularity and rule of thumb, not a necessary result. A conceptually necessary, unavoidable gradient in the direction of expression is plausible but debatable. Betrayal by means of a micro expression may simply never happen, whether due to rehearsal of alternative display behavior, pushing back the expression of one emotion that thereby remains incomplete

*via* substituting the expression of another (e.g., sadness for would-be shame), insufficient strength of the candidate emotion. But it often does.

The suggestion for further empirical inquiry is that a person's empathic receptivity can be trained to further attunement to these micro expressions in spite of their not being visible to the unaided eye (without slowing down the video replay). The suggestion is that a significant "gray area" is available that represents a zone of proximal development of empathic receptivity. Some empathic individuals may already be good at it and get better with training. Alternatively, some individuals may tend to wear their "hearts on their sleeves," meaning that their micro expressions have less granularity and are therefore more visible without the aid of slow motion or other technology.

While individuals have control over a wide range of facial expressions and this range can be extended with practice, some, perhaps many, muscles in the face are independent of an individual's voluntary control. These muscles are "hard-wired" to the nervous system and form micro expressions that take less than 0.2 seconds to discharge. For example, some of the muscles around the eye are only activated by a smile that recruits real enjoyment as opposed to turning one's lips up into a "grin and bear it expression" as reported by Paul Ekman, whose research broke new ground around micro expressions:

> Our research confirmed Duchenne's assertion that no one can voluntarily contract the *orbicularis oculi* muscle (it "does not obey the will"), although it is only part of that muscle that is hard to contract voluntarily.[34]

[34] P. Ekman. (2003). *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*, New York: Henry Holt, 2003: 206. [Editorial note: This example is also discussed in Chapter ___ The Significance of Neurology for Empathy: The Philosopher's Cerebroscope.

A similar consideration applies to the red margin of the lips that become narrower in anger as the lips become thinner, which is a very hard action to inhibit. The rate of blinking increases when an individual is nervous. Eyes widen with excitement. Nostrils flare with excitement. Such micro expressions are fleeting and highly transitory. In everyday situations they can reveal a person's underlying emotion, ambivalence, or mixed emotions. For instance, the would-be terrorist has nothing but contempt for those he is about to blow up and a brief expression of it flickers across his face.[35]

In another example, an individual contemplating suicide but feigning happiness in order to get a weekend pass home from the psychiatric ward was video taped as part of an unrelated training exercise:

> …[I]nspecting each gesture and expression in slow-motion to uncover any possible clues to deceit. In a moment's pause before replying to her doctor's question about her plans for the future, we saw in slow-motion a fleeting facial expression of despair, so quick that we had missed seeing it the first few times we examined the film. Once we had the idea that concealed feelings might be evident in these very brief micro expressions, we searched and found many more, typically covered in an instant by a smile (Ekman 1985: 17).

After this review, she acknowledges being desperately unhappy and wanting to escape to commit suicide. The tape was literally slowed down in order to reveal what was present but did not reach the threshold of conscious awareness. In some instances, this explains the "hunch," "intuition," or "gut feeling" that causes doctors (or police) to engage an individual and ask a few innocent questions. The trace of sadness or contempt would be in the system of the perceiving individual and available for empathic receptivity; but

[35] For example see "Silent signals: security concerns bring new focus on body language: FBI takes 'science' more seriously," *The Wall Street Journal*, August 15, 2002, page 1.

would be on the borderline of awareness. If one deploys additional attentional resources, the fine-grained micro expression may be identified and recruited to enrich one's empathy.

Granted that the human eye cannot interpret more than about 16 frames per second, if you slow down the film, the viewer can penetrate up to an order of magnitude in the direction of fine-grained micro expressions. This is a reversal of our contingency. It motivates ("explains") the benefit of studying art and music for increasing our empathy. Improvements through practice in fine-grained sensory and affective discriminations in the realm of humanlike artifacts—works of art—are readily transferable to the domain of human empathy.

Empirical research and training in introspective methods can develop and enhance this attention to detail—this empathic receptivity to a trace or signal in the form of a micro expression that betrays an otherwise unexpressed emotion. Training and practice can improve the result, though there is likely a threshold below which training will not make a difference. Finally, there is no easy answer to someone's deception when employing the Stanislavsky Method of acting to cause a smile based on genuine recollections of happy times in the past or when sincerely fooling oneself. Interpretation remains an essential part of the empathic process and the indeterminacy of translation is inevitable—as is anomalous monism--when one allows that the cause and target of the emotion is part of the context. This best way to get someone to else to accept a lie – or in the case of the

theatre to willingly suspend disbelief in the pretense – is paradoxically for the liar to believe the story himself.

### Key Case Two: Subtype Ambivalence

The second subtype is the example of becoming aware of the ambivalence of the unexpressed emotion in the process of bringing to expression that something = x that is not nothing. The "x" turns out to be a thicket of multiple, mixed emotions. For example, something happens that arouses an emotional reaction—a person is the object of a joke—and he does not know whether to be happy (and laugh) or angry. He suspects he was just insulted, and perhaps he really was. The joke was about his profession. He is a lawyer and the punch line is "The shark didn't eat the attorney. Professional courtesy!" The unexpressed emotion (the affective component will be prominent) is given as an immediate reaction before the person expresses it. But in expressing the unexpressed emotion the person will confer an individuality on it, which opens and transforms an opaque and vague something = x into a definitely accessible feeling. The person is now aware of the complex emotional reaction in a way that was previously not the case. The person laughs as a way of dispelling the reaction, whether anger or shame or sadness, is not material at this point (e.g., M. Lewis 1991: 130). A person can end a laugh with a snarl, though it does not show the person at his best. Many expressions of emotions are "compromise formations," that express ambivalence in "grin and bear it" smiles or other mixtures of happy and sad, angry and afraid. An impedance mismatch occurs between the sequential output device (the physiognomy-physiology) and the multimodal milieu of ambivalence. The person laughs and a micro expression of hostility passes over the face

before being masked in a polite smile. “Ha, ha, ha, [Grrr]!” This binds the vague and ambiguous reaction to a particular facial and physiological expression and completes at least one of the emotions—but the other one remains incomplete and that, under this interpretation, can be the start of trouble. Empathic receptivity taps into the contingent but powerful tendency in the unexpressed component to break out and be known for what it is. Empathy is particularly powerfully constellated by ambivalence, homing in on the component that is betrayed, not displayed.

Social referencing (Hobson 2005) has a similar profile to ambivalence and is quite remarkable. In the case of small children, it shows one form of ambivalence emerging. A toddler will be walking along and fall over suddenly, then look towards its mom or dad to see if the care-taker looks worried or happy. If the care-taker is smiling, then the child will laugh; if the care-taker looks worried, then the child will cry. This is a crisp example of the child’s referencing the other to check how serious mom considers the fall prior to expressing any emotion in the matter. The child is checking with mom to see how he should feel. The latter feeling is evidently still an unexpressed something = x where the care-taker has a critical role in deciding whether to bind the “x” to fear (“hurt”) or bind the “x” to happiness (“having fun”). In a marvelous example of emergent empathy, the empathic care-taker expresses the emotion on behalf of the child, which emotion is then taken up and further processed and expressed by the child in reciprocal affectivity and attunement with the care-taker’s response, completing the circle. I suggest that something analogous occurs between adults when, after giving an account of some dramatic vignette, the one person turns to the other with a significant pause in order for the other to

respond with a demonstration of having gotten what the first one went through. We turn to the other to find out whether what we experienced was fun or fearful, a source of hope or despair, etc., and use the spontaneous response of the other to guide how we really feel about what occurred. This does not rule out that one person is often looking in advance for a particular reaction and to "get a rise" out of the other; the adult scenario is more complex than that with the toddler; and the "get a rise" is not necessarily what empathic understanding based on a gracious receptivity is going to provide her or him.

This inquiry into the fine-grained distinction of empathy may lead some to object: Is empathy really required to see that another is angry (say) when the brothers pound Joseph on the head? In the final analysis, yes, empathy is indeed required to recognize a gross instance of anger; but it may seem otherwise since empathy is not properly put to the test by such an instance. [36] Another instance with finer details - e.g., ambivalence - is required before empathy can be activated in any interesting way and put to the test. For example, if Dan senses the sadness of the other brothers at the loss of their father's love whereas everyone else only notices the great anger, then a distinction only possible through a greater ability to make fine-grained distinctions of Dan's empathic receptivity is made evident. While all the brothers have the same sensations—hear the swear words, terms of abuse, see pounding of fists—only Dan (in this case) is aware of the others' sadness. Dan is receptive to something to which the others are indifferent and unaware. His empathic receptivity is more fine-grained - one might say "delicate" - than the others.

[36] This example is also considered in Chapter ___ on The Significance of Neurology for Empathy: The Philosopher's Cerebroscope.

### Key Case Two: Subtype Catching Oneself in the Act

A third subtype is a variation on the second. An individual becomes aware that he was in a particular emotional state without it being available to his attention, but when his attention was called to it (the emotion), then he recognizes and acknowledges the emotion whether or not he further expresses it. Consider: A person is given a beeper and agrees to writing down in a notebook the answer to one question (phrased in various ways) when the beeper goes off at any arbitrary time: "What was occurring in your inner experience at the moment of the beep?" One individual discovered that he had frequent angry thoughts about his children. The individual had been unaware of those thoughts prior to this experiment in descriptive experience sampling (DES).[37] This "angry-towards-children" affect was readily available, but as a matter of contingent circumstances, the individual was unaware of it. The affect was descriptively unconsciousness, pre-conscious, but readily acknowledged when the beeper went off, causing the angry thoughts and feelings to become attentionally available to awareness. This is in contrast to the situation with Freud's *Fehlleistungen* or the couple in marital therapy where they cannot acknowledge the emotion they experience for one another. In the latter case, the contempt is betrayed in micro expressions, visible when the tape is slowed down, but that is not all. The contempt is also expressed, not merely displayed, by being displaced onto other objects—the husband is less cautious in explicitly expressing contempt for the wife's pet dog, saying that it smells bad. The husband's contempt for the wife remains unexpressed, and the marriage is (probably) on the rocks. Nor does this brief vignette

[37] R.T. Hurlburt and C. L. Heavey, (2001), "Telling what we know: describing inner experience" (2001), *Trends in Cognitive Sciences* 5: 400-403 cited in A. I. Goldman, *Simulating Minds*. Oxford: Oxford University Press, 2006: 242. [Editorial note: This experiment is also discussed from a different perspective in Chapter ___ on Empathy and Introspection.]

imply that expressing the contempt would have saved the marriage. Further counseling is required.

The example of the individual who is ambivalent about the joke but does not express or acknowledge it even to himself, weighing the anger and amusement, masking the anger with a mild smile according to a social display rule (say, because of peer pressure to maintain group harmony) is an intermediate case. If the individual would deny being angry even off-line, in private with the experimenter, then the attentional availability (of the anger) is strongly disassociated due to social pressure. It is not available to attention. However, he might go home and kick the dog—that is, take the incomplete emotion and displace it onto an unrelated object in order to try to complete it. Unfortunately, this dog is in for a rough time, since, under this scenario, the kick is a betrayal of the emotion, not a display of it, and what is betrayed remains a something = x that is not nothing but not quite something either. Empathy shows it will happen again-and-again until completed.

In contrast, the father's anger in the DES experiment does get expressed ("bound") by being written down in the experimental log. The beeper goes off. His attention focuses on what he was just feeling and thinking. He has no problem acknowledging it. It is not blocked by social pressure, a display rule, or some version of Freud's dynamic unconscious. This example collapses into the example of becoming aware of the emotion in the process of expressing it if he does indeed express it; or into key case one if he does not. Had this individual not participated in the DES experiment, he might never have gotten around to providing his children with the appropriate parental coaching. Of course,

the risk is that, if left incomplete, the unexpressed emotion will behave as a free floating, unsaturated, unsatisfied emotion that will bind itself opportunistically to various unrelated issues, damaging the relationship bit-by-bit.

### *Key Case Three: A Limiting Case*

The third key case initially resembles key case number two where obstacles get in the way of recognizing the unexpressed emotion, and the unexpressed emotion is unwittingly betrayed. Recall Ryle's example of vanity and the vain person who never *feels* vain. This is a limiting case in that the allegedly and contingently unexpressed (emotional) something = x to be surfaced is not surfaced. The feeling of vanity cannot be surfaced since it does not exist. The vain man experiences sadness when he is not acknowledged for his exaggeratedly self-centered behavior and pride when he is. He experiences a full spectrum of emotions such as happiness, sadness, pride, disappointment; but it is just that there is no specific feeling characteristic of vanity. This leads to the issue of what it means for vanity to be displayed or betrayed. In the case of vanity a social norm becomes a source of resistance to the vain person acknowledging vanity. That obstacle is: vanity is not something that is a source of positive recognition. Calling someone "vain" is a criticism. The fancy clothes and self-congratulations of the vain individual are meant to display how special he is and worthy of honor, not confess the short-coming of vanity.

In the case of vanity, the distinction between displaying an unexpressed emotion and betraying it collapses. Vanity does not try and hide. It presents itself front-and-center, often in the form of ostentatious dress, but also in boasting about accomplishments never

quite completed (at least not by oneself). It is always on display. When it stops being displayed, it falls back to a something = x that is not a feeling. The affective aspect is missing unlike most of the basic emotions such as fear, sadness, anger, surprise, contempt, etc. that, at least initially, have a strong felt component. In itself, this arouses one's empathic receptivity, simply because it (the nonexistent feeling of vanity = x) is conspicuous by its absence. This doesn't mean that it has to be there and is hiding. It might (and does) mean that emotions are themselves ambiguous and not a natural kind, occurring as irruptive emotional reactions, social roles ("pretenses"), or even moral sentiments.[38]

If the vain man were to describe the sensations, affects, and "physical agitations" (as Ryle might say) that accompany the implicature that he is the center of the conversation, then the vain person would say that his heart swelled with a feeling of pride upon being acknowledged and recognized, that he felt happy to make a contribution, and that he looked forward to even more success. He would react as if people were recognizing and honoring him, as the guest of honor, even though others were just being polite or mildly amused at the ostentatious display.

The vain person's acquaintances would say that he acted so as to gain the attention, admiration, appreciation, and even envy of others. It is not so much that there is a "particular palpitation or pricking of vanity" (indeed there is none), but rather that the vain man regards (understands) his sensations and behavior (with its interpersonal

[38] P. E. Griffiths. (1997). *What Emotions Really Are*. Chicago: University of Chicago Press, 1997.

context) as pride or self-centered admiration and the majority of the rest of us describe it – that is, interpret it - as vanity. The vain man asserts that the song is written to honor him whereas it is written to make fun of him.

The emotion of vanity is a social role that includes a form of self-deception, but not about what one is feeling, rather about the role one is playing. When we go in search of feelings, using empathic receptivity, then there is no special feeling of vanity, but rather the usual feelings of happiness and pride, granted that these feelings may not be appropriate. The vain person feels "justified" pride; we feel boredom at his foolishness. The vain person feels "honored" at his success; we feel impatience at his arrogance. The vain person feels "validated" in his exceptional worth; we are annoyed. The distinction between display and betrayal collapses in the case of vanity, or, alternatively, the display of vanity is precisely its betrayal. The vain person has nothing to hide. Thus, the vain person has nothing to betray.

Of course, one may redescribe the situation and offer the objection: the vain person does indeed have something to hide, namely, that he or she is presenting a facade behind which there is nothing substantive by way of character, integrity, or depth of feeling. The vain person is hiding his or her hollowness. When we become aware of the other's vanity that is the distinction that is grasped – a fancy presentation ("outside") masking a lack of depth ("inside"). When the vain person becomes aware of precisely this distinction – I put on a good show but it is all show – all cowboy hat and no cattle, as they say in Texas about "urban cowboys" – the vain individual has ceased to be vain at least for the

moment. So it is the case after all that the distinction between display and betray collapses.

Significant work is required to imagine a scenario where vanity might be betrayed rather than displayed, and even then it will seem anomalous. Vanity is supposed to be an emotion, and it is indeed that - and more. Suppose an organization existed, "Vanity Anonymous," whose members were individuals trying to overcome their vanity. As a vain person, I go to a meeting and confess, "I am a vain." The club members would meet and share tips and techniques on how to give honest and sincere appreciation to others, dress and speak modestly, and recognize the accomplishments of others ahead of their own. But then one individual slipped up and went on a tear, taking credit for the deeds of others and acting as if he were the center of attention. Once again, he is so vain that he hears the song and thinks it is about him. Upon realizing that he had failed to live up to his goal of overcoming his vanity, he will say, "I messed up. I lost control. I betrayed my vanity. Let me renew my commitment to modesty." Yes, a distinction opens up here between vanity and its betrayal. Yet what was betrayed was a character trait of vanity that was in the process of being reworked through practice, not some occurrent and contingently unexpressed emotion.

Vanity is precisely one of those emotions where little or no agreement occurs between self-description and the way others describe the individual (namely, as a vain person). In the case of basic emotions or irruptive social emotions (e.g., righteous indignation), we may debate about whether a person is angry or sad, etc., and sometimes conclude, "Yes,

his anger was palpable; it could be felt." However, the vain person's vanity is never felt. Never. The observers feel boredom or frustration or annoyance or eventually anger. Indeed if the vain man were to recognize what he was feeling, then his prideful but little justified honoring of himself would turn to shame. If that recognition occurred, he would no longer be vain. He would see himself as others see him; and, in this case, the sight is not a pretty one, resulting in a negative evaluation of himself.

Given the collapse of the display/betrayal distinction, vanity as an unexpressed emotion starts to look either like a disposition (which is why Ryle seized on it as a useful case) or like a character trait, which has many dispositional aspects. In either case vanity as an unexpressed emotion vanishes into a cloud of behavior. To be is to be perceived - to be (vain) is to be perceived (as vain).

Ryle's assertion "vain man never feels vain" was a brilliant slight-of-hand. Nor does it follow that as a consequence of the vain man never feeling vain that "the angry man never feels angry." He does. The angry man feels angry. Likewise, the sad man feels sad and cycling through all the basic emotions discussed by Darwin and Ekman. The sad man does indeed feel sad; and unless this particular man is making excuses of the kind that would cause us to say he is in denial – "I'm not sad, just tired" – the sad man will readily acknowledge it. "I just can't get a date – that's what I'm sad about." "Mom always liked my brother Tom best. That's the source of sadness in my life."

However, saying "He is angry, but he does not feel angry" is not a logical contradiction. A person who speaks that way knows the meaning of the word "angry," even if such occurrences are the exception not the rule. Saying "He is angry, but he does not behave angrily" is not a logical contradiction. A person who speaks that way knows the meaning of the word "angry," even if such occurrences where people are self-controlled are the exception not the rule. But what about "He is vain, but he does not behave vainly"? Well, then it depends. Perhaps he was vain, but overcame the character trait through long practice. Perhaps he will slip up under stress—no contradiction. But if the individual never behaved vainly, then a logical contradiction is in the offing. If he never behaves vainly, then the person who talks that way does *not* know the meaning of "vanity." Both the expression (and lack of it) of the emotion of vanity - and the vain man who never behaves vainly - evaporate into a puff of behavior. This is *not* an exception, though it is a limit case. (Arguably it would be yet another exception among many if one were trying to provide a unified definition of the emotions, which, however, is not the case here.)

Thus, vanity is a paradigm of the emotions as social roles. The twelve angry men in the jury room in the film of the same name were also in an occurrent emotional state characterized by a social role. Though they were not aware of it as a matter of explicit attentional focus, they were in the role of being angry about urban decay and related urban violence in the 1950s, especially gang violence with switch-blades, and this individual, the accused, stumbled into their drama. The twelve jurors were about to convict an innocent individual of a heinous crime. Their self-deception was palatable. All but one of them "knew" that the accused was guilty. The accused fit the stereotype. He fit

the role. He fit the profile. So did eleven of the twelve jurors. The circumstantial evidence was compelling. Only one individual knew that he did not know. This opened up his empathy, in this case not only with the accused and with his fellow jurors, but with gang members' use of switch-blade knives. Based on his empathic interpretation of the role of innocent Latino mistaken for a member of a gang – and the distinction between stabbing upwards versus downwards with the switch blade - the lone hold out was able to convince his fellow jurors that the accused was falsely charged and indeed innocent.

## *Emotions are Storied*

Emotions are inherently dramatic. They are the occasion for tears, hostile explicatives, flight in the face of the threatening, and happy endings where the hero gets her man and rides off into the sunset. It is not surprising that we find emotions and emotional behavior at the core of dramatic representations from Greek tragedy to modern Harlequin romances. Even if we are hard-wired, as Ekman (and Darwin) would argue, by our evolutionary heritage to cry in sadness and clench teeth in anger, we are motivated by powerful cultural and societal constraints to try to make sense out of our emotions by storying them. In these stories emotions become both causes and reasons, depending on context, for human actions and reactions. The emotions are both drivers of the action and the target of explanation through the action of the narrative.

It is likely that the basic emotions arise out of the undifferentiated biological substratum that we share with our simian and canine relatives. They arise out of it and never completely leave it behind. But as humans we quickly envelope them in a network of

semantic distinctions that makes the emotions accessible, helps us cope with them, control them if they require controlling, and makes them tractable, even if only temporarily and intermittently.

Saying that the emotions are storied means that the emotions are a source of stories as they follow their inherent trajectory – part physiological, part intellectual – to gain expression. In another example from literature that without expression, emotions are incomplete in Gabriel Marquez's *Love in the Time of Cholera*. The suitor spends a whole life secretly in love with the lady across the street - secret because it was not communicated in a way that it was ever empathically received, even though the marriage contract was discussed - is particularly poignant, because the mere expression of the love is necessary for its completion even if the love is not reciprocated. The expression of the love is not necessary for its existence, for even unexpressed it has the shadowy life of a wheel that moves no part of the interpersonal world and represents an inner barrier and obstacle to the lover's participation in life. But the expression of love is necessary for its completion. Only if finally expressed – and rejected or disappointed – the would be lover is able to love again and find another – or if returned, explore a new life with the beloved where after ten years of marriage and several children, the meaning of "love" will have been transformed beyond recognition from the perspective of the infatuated suitor. This does not means the emotions are judgments or should be expressed as such. Rather that

the emotions are storied means that a wealth of narrative surrounds them and they will exhaust all the narrative we bring to them.[39]

## *Emotions Unbound*

The denial that there are unexpressed feelings is a capitulation to that shallow kind of behaviorism which tries to reduce feelings to their expression. No such capitulation has occurred here. The concept of emotion itself is a hybrid, as some might say, "mongrel," category. It encompasses hard-wired, irruptive reactions as well as social roles ("pretenses") such as vanity, road rage, the behavior expressed by and in righteous indignation, films such as *The Rebel Without a Cause*, *Twelve Angry Men*, "martyrdom" by means of a suicide bomb, the shame of a woman scorned with Hester Prynne's scarlet letter, as well as "wild pig" behavior and ghost possession in certain Polynesian cultures.[40] An entire class of emotions requiring social referencing to others such as shame and guilt and including moral sentiments contributes to solving the problem of how creatures of limited generosity and strong self-interest can reasonably adhere to ethical commitments.[41] However, as regards the dynamics of such emotions, they are all subject to the individualization that occurs in binding an inchoate, emergent something that is not nothing = x to specific forms of expression. Without expression, emotions are incomplete.

---

[39] After I completed writing this material, a colleague pointed out to me that Peter Goldie has proposed a similar approach, which I hereby acknowledge: Peter Goldie. (2000). *The Emotions: A Philosophical Exploration*, Oxford, UK: Clarendon Press. My approach to narrative is based on Ricoeur (1985), though he does not apply it to the emotions or empathy as such.

[40] Ian Hacking. (1999). *The Social Construction of What?* Cambridge, MA: Harvard University Press, 1999.

[41] R. H. Frank. (1988). *Passions within Reason: The Strategic Role of the Emotions*. New York: W. W. Norton. On "social referencing" see Peter Hobson. (2005). "What puts the jointness into joint attention?" in *Joint Attention: Communication and Other Minds: Issues in Philosophy and Psychology*, eds., N. Eilan, C. Hoerl, T. McCormack, and J. Roessler. Oxford: Clarendon Press, 2005.

The importance of letting our conceptual analysis be informed by empirical research is evident (for example) in that micro expressions were unappreciated prior to the Ekman's research, in which he tried to refute Darwin and failed productively and insightfully. In every case, expression does not just describe emotions, it individualizes them, and brings them forth by binding an unexpressed something = x to affective, physiological, and semantic distinctions. Without being expressed, the emotion of which the feelings and sensations are components is incomplete.

Unexpressed emotions *live*. Empathic receptivity is particularly affected by the betrayal of emotions that are contingently unexpressed, sincerely denied, ambivalent, emergent, and appear as social roles ("pretenses"). Alternatively, working from explicit behavior towards the unexpressed something = x, empathy enables a person to put himself empathically ("as if") in the position of the other, who is engaged by an emotion as a social role such as vanity. Thus, empathy goes in both directions. The "as if" approach itself can arouse empathic receptivity to fine-grained distinctions at the threshold of awareness. This is an area of proximal development for empathy receptivity and empathic understanding in completing the disclosure of emotions, expressed and unexpressed.