

This article was downloaded by: [Princeton University]

On: 24 February 2013, At: 11:50

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Cognition & Emotion

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/pcem20>

Basic emotions, rationality, and folk theory

P. N. Johnson-laird ^a & Keith Oatley ^b

^a Department of Psychology, Princeton University, USA

^b Applied Cognitive Science, Ontario Institute for Studies in Education, Toronto, Canada

Version of record first published: 07 Jan 2008.

To cite this article: P. N. Johnson-laird & Keith Oatley (1992): Basic emotions, rationality, and folk theory, *Cognition & Emotion*, 6:3-4, 201-223

To link to this article: <http://dx.doi.org/10.1080/02699939208411069>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Basic Emotions, Rationality, and Folk Theory

P.N. Johnson-Laird

Department of Psychology, Princeton University, U.S.A.

Keith Oatley

*Centre for Applied Cognitive Science, Ontario Institute for Studies in
Education, Toronto, Canada*

Answering the question of whether there are basic emotions requires considering the functions of emotions. We propose that just a few emotions are basic and that they have functions in managing action. When no fully rational solution is available for a problem of action, a basic emotion functions to prompt us in a direction that is better than a random choice. We contrast this kind of theory with a componential approach which we argue is either a version of the theory of basic emotions or else leads to the doctrine that emotions are mistaken tenets of folk psychology. We defend the psychological reality of the folk theory of emotions, and we argue that universal basic emotions make it possible to understand people from distant cultures, and to translate emotional terminology from one language to another. Finally, we show how theories of basic emotions can be tested, and indicate the kinds of empirical result that can bear on the issue.

INTRODUCTION

How many emotions are there? There are several possible replies to this question. One is that the question is meaningless, but this response amounts to rejecting the ordinary concept of emotion. It is akin to answering that emotions do not exist—that they are false tenets of folk theories, i.e. of the common sense theories that lay people have about mind and behaviour, which contrast with scientific and other kinds of specialist theories (D'Andrade, 1987). Another answer is that although individuals experience only a finite number of emotions in their lifetimes, there are indefinitely many possible emotions that they might experience. This view can be coupled with the claim that one never experiences the

Requests for reprints should be sent to Dr P.N. Johnson-Laird, Department of Psychology, Princeton University, Princeton, NJ 08544, U.S.A.

same emotion twice. If there is an indefinite number of emotions, then we can ask how they relate, if at all, to one another. Each emotion might be unique and unrelated to any other. Such a conception is metaphysically defensible but removes emotions from the domain of scientific investigation. Each wave on the seashore is unique, but the science of hydrodynamics idealises waves, and from this abstract standpoint treats waves that are, in fact, distinguishable as the same.

In this paper we will put a different view, that folk psychology and scientific psychology both have something in common. Folk psychology treats different experiences as instances of the same emotion, so in daily life, we talk of different occasions of fear, anger, happiness, and so on. Likewise, a science of emotions is bound to treat different emotional experiences as members of the same class. Hence, we can ask our question again, but slightly differently. How many sorts of emotion are there?

One answer is that there is a small finite set of distinguishable emotions that are the bases of all emotional experiences. This is the hypothesis of basic emotions as primitive unanalysable elements at the psychological level of the system. Another answer is that there are many sorts of emotion, but every distinct sort is generated from among the same finite set of components, much as, say, each chemical molecule is constructed from atoms. Hence, all emotions would be on a par—there would be no sense in which some emotions were basic whereas others were complex.

To be more specific, a theory of basic emotions might analyse embarrassment as founded on the basic emotion of fear, with some other non-emotion component such as a cognition of a particular sort, i.e. knowing that one is the object of unwelcome attention. A componential theory in which no emotions are basic might distinguish embarrassment from fear in terms of different appraisal-response elements that make up each kind of emotion. Such a theory then holds that because of its components each type of emotion has its own unique psychology and physiology, although perhaps with some components in common. Both the theory of basic emotions and the componential theory postulate that an emotional experience depends on various elements. The crucial distinction is that no components can be an emotion *per se* in the componential approach, whereas one component of any emotional experience is always a basic emotion according to the basic theory. Both theories assume that a science of emotions is possible only if there is a finite basis for emotional experience. Both assume that emotions can be taken to pieces analytically. The question is: Is there always a piece that consists in an irreducible basic emotion or are there sub-emotional elements, such as appraisal-response components, into which all emotions can be analysed?

There are various ways in principle of establishing the nature of the finite basis of emotions. In this paper, we will examine the lesson that the function of emotions can teach us with reference to limitations of human

rationality. Next, we will enquire into the set of basic emotions. Once we have outlined a theory of them, we will contrast it with a componential theory, and we will show that the componential approach turns out to be a version of the theory of basic emotions or else leads to the doctrine that emotions are a false tenet of folk psychology. We will defend the psychological reality of the folk theory of the emotions, and we will counter arguments based on the difficulty of translating emotional terminology from one language to another. Finally, we will show how the theory of basic emotions can be corroborated, and describe some of these results.

THE NEGLECT OF FUNCTION

What function, if any, do emotions serve? The question has been somewhat neglected by theorists, and one can read much on the cognitive underpinnings of emotions that does not address this issue. Yet, in our view, it is the key to whether or not there is a small set of basic emotions: The hypothesis of basic emotions makes sense only if it elucidates problems faced by the cognitive system. Although elements of current theories of emotions can be traced back at least to the nineteenth century, the intellectual history of the topic has lacked cumulative coherence. Without any clear sense of the psychological function of emotions, it has been difficult to generate more than a patchwork of ideas and observations. With such a sense, perhaps the scientific understanding of emotions will become cumulative, analogously, say, to the understanding of visual perception.

One reason for the neglect of function is probably the pervasive influence of William James. Like Descartes (1649/1911) he classed emotions with perceptions (e.g. James, 1890): They are perceptions of events inside the body. Beyond his claim that not all emotions are accidental, he had little to say about what purposes they might have. Because, according to James, emotions are percepts of bodily feedback from physiological changes, or from actions that have already taken place, emotions occur too late to affect either the control of these actions or the decisions that led to them. Just as there are indefinitely many percepts of the outer world there are, in James's view, indefinitely many feelings, each reflecting a particular pattern of proprioception and physiological perturbation.

James's influence may account for the importance that many theorists place on emotions as primarily pleasant or unpleasant, that is, as "valenced" (e.g. Frijda, 1987; Ortony & Clore, 1989). If emotions are not a system for the immediate control of actions, then they are important end-points giving colour to experience. Incidentally they can then have motivational consequences. People strive to attain them if pleasant, and to avoid them if unpleasant (Hammond, 1990).

In short, James's theory of emotion contributes to the development of a powerful tradition. His conclusions, perhaps serendipitously, resonate with the high valuation of "experience" as such in Western culture. Pleasantness and unpleasantness have become the crucial characteristics of emotions over and above their own strict individuality. Within this tradition, the possible existence of a set of basic emotions seems both unattractive theoretically and intractable empirically.

RATIONALITY AND THE FUNCTION OF EMOTIONS

Analyses of mental processes in cognitive science assume that each process has functions independently of its particular embodiment. From this perspective, emotions should have a function that could be embodied in a system based either on carbon-like humans or on silicon-like computers. As many people within cognitive science have argued, function is accordingly best thought of in terms of the design of the system. During natural selection systems are fitted to functions, although as a means of design, evolution is notoriously a "tinkerer" not a grand architect. A priori there are many possible designs to enable organisms to cope with their environment.

The simplest possible design relies on "fixed action patterns" and is found in insects. Consider, for instance, the common tick, which is a parasite of mammals. According to von Uexkull (1957), the female tick lacks eyes, but at one stage in her life cycle the photosensitivity of her skin triggers the action of climbing a bush from which she then hangs. She lets go only when a second trigger occurs: The smell of butyric acid, which is secreted by the sweat glands of all mammals. If she happens to land on a passing animal, a third trigger comes into operation: The warmth of her host's body. Propelled by this taxis, she burrows through the hair to the skin, and there she punctures the skin and fills herself with blood. Once full, she drops off the animal to lay her eggs on the earth. And the cycle continues anew.

This sort of design works well when classes of events can be mapped one-to-one on to appropriate responses. Perfect performance is impossible, e.g. a tick may drop but miss the passing mammal. Yet, the solution is rational in the following sense: All that is necessary for a reasonable chance for individuals to survive and to reproduce is built into the species-specific procedures for action. In principle, there are no uncertainties about what to do: The stimulus either unlocks the fixed action pattern, or not. Of course, this certainty can be the undoing of a species if there is a significant change in its environment.

At the other extreme in the theoretical series of designs are those that are impeccably rational. They are maximally flexible because they enable

the organism to determine which goals to pursue at any point in time, and to decide at each choice point the best course of action in pursuit of those goals. No contingency is unanticipated, and performance is invariably optimal. Creators of artificial intelligences have aspired to such designs; philosophers have argued that they are realised in human thought (e.g. Dennett, 1978, p. 20; Cohen, 1981); and psychologists have claimed that apparent errors are merely failures in performance that do not impeach the underlying rationality of the system (e.g. Henle, 1978).

In designs based on fixed action patterns or on impeccable rationality, there is no occasion for anything corresponding to an emotion. There are no surprises, no misunderstandings, no irresolvable conflicts. Human beings are neither equipped with a set of responses each matched to an important stimulus, nor do they possess impeccable rationality. A fully rational system of thought is a paragon that cannot be realised by any finite device. Any set of observations is compatible with an infinitude of different valid conclusions, and so no finite organism can follow up all of them (Johnson-Laird, 1983; Cherniak, 1986; Stich, 1990). Moreover, human reasoners make genuine mistakes in reasoning—mistakes that they even acknowledge in some cases. They make invalid inferences that should not occur if their thinking were guided by valid formal rules of inference (Johnson-Laird & Byrne, 1991). In short, to paraphrase de Sousa (1987): Human beings are neither insects nor omniscient, omnipotent gods.

If impeccable rationality is impossible, what design is embodied in human beings? Johnson-Laird and Byrne (in press) argue for a significant modification of the competence-performance distinction. The original distinction hinged on the idea that rational competence is based on valid rules of inference, which, like the rules of grammar, might sometimes be inadequately reflected in actual performance. The new notion of rational competence depends instead on a meta-principle: An inference is valid provided that there is no model of the premises in which its conclusion is false. Individuals have a tacit grasp of this meta-principle, and they put it into practice by building mental models of premises, drawing useful conclusions from them, and then searching for alternative models that might refute such conclusion. But they have no grasp of any specific logical rules, and they have no comprehensive algorithm for valid thinking, i.e. for searching for models that refute conclusions. The meta-principle is defensible as a rational requirement for any system for deductive inference, although it alone does not guarantee the validity of inferences. To argue that errors arise as result of performance factors is misleading, however, because it suggests a failure to put into practice correct rules, whereas there are no rules to put into practice, only the higher-order meta-principle. This principle is compatible with the observations of deductive failure, and with the arguments against impeccable rationality.

Granted that reasoning is fallible and time-consuming, Oatley and Johnson-Laird (1987), following Simon (1967), and a tradition of cognitive theorists, proposed that the function of emotions is to fill the gap between fixed action patterns and impeccable rationality. For many species, including *homo sapiens*, the world is too complex to form perfect mental models, so events and the outcomes of actions are often unanticipated. The problem that Simon identified is that complex systems acting in the natural world, as opposed to a simplified microworld, need something equivalent to interrupt signals in computation. Such signals are necessary in systems that have limited resources, and that need to be influenced by unforeseen events demanding urgent attention. Emotions, as Simon noted, seem to be co-extensive with the occurrence of such problems. They arise particularly when individuals have many concurrent goals, including mutually incompatible ones, and their resources of time, ability, and processing power, are too limited to make a fully rational choice. Moreover, social mammals often cannot achieve their more valuable objectives alone, and so they need to interact with others. Co-operation calls for mutual plans, but it is impossible to guarantee that copies of the plan kept by each partner are identical. Competition calls for antagonistic plans, and it is impossible to determine their outcome. The biological system of emotions offers a solution to these problems, particularly those that arise from the limits of rational principles to govern or to predict complex social interactions. Emotions enable social species to co-ordinate their behaviour, to respond to emergencies, to prioritise goals, to prepare for appropriate actions, and to make progress towards goals. They do so even though individuals have only limited abilities to cogitate.

Emotions guide individual and group behaviour. Social mammals are unable to determine the best course of action at many of the junctures in their lives. Even in humans, the resources for rational thought are often too slow and too error-prone to solve this problem. The function of emotions is accordingly to bridge the gaps of rationality. We argue that this bridge is possible only if many specific junctures can be mapped into a few broad classes of reaction.

We have proposed (Oatley & Johnson-Laird, 1987) that the cognitive evaluation of a juncture in action calls into readiness a small and distinctive suite of action plans that has been selected as appropriate to it. Each basic emotion thus prompts both the individual and the group in a way that in the course of evolution has been more successful than alternative kinds of prompting in broadly defined, recurring circumstances that are relevant to goals. Thus, when the broad class of event occurs that indicates achievement of a subgoal that increases the probability of attaining a goal, then its cognitive evaluation initiates an internal emotional signal. We propose that emotion signals of this kind have no propositional content or syntactic

structure: They have a control function rather than an informational function. The signal that is sent when subgoals are achieved acts to prompt the individual to continue the same line of action. When a goal is lost, a different emotion signal is sent. It prompts the individual to disengage from that goal. The internal emotional signals have causal effects within the organism, preparing it physiologically for each general class of actions. In the case of human beings, the signals can in addition be experienced subjectively as emotions. The signal caused by a successful achievement is experienced as happiness, and the signal caused by the loss of a goal as sadness. An important consequence of the ensuing actions is the communication of the individual's emotional state to others in the same social group—an example is the distinctive type of alarm signal sent by certain social mammals and birds. The receipt of such external signals has emotional consequences for these other individuals too.

If the emotional guidance of action is to be rapid, successful, and independent of reasoning which is too time-consuming, then the cognitive evaluations must be coarse and the resulting suites of actions must be broad and flexible. There are two key issues here. First, many events in the world must be mapped on to a relatively small number of categories, which each elicit a distinct set of bodily, behavioural, and (at least in the case of humans) phenomenological consequences. If there were very many categories, then the problem of deciding amongst them would re-emerge as a time-consuming matter. Secondly, the small repertoire of actions triggered by a particular emotion must be useful to a wide class of specific triggering events. For example, if there is a conflict in goals because an event threatens an individual's safety during the course of another action, then the emotion of fear prepares a small repertoire of actions, which includes stopping the current action, checking everything that has been done recently, monitoring the environment, fleeing, being prepared for fighting, physical exertion, or bodily harm. In the case of human beings, the repertoire can be supplemented with action sequences that have been practised. The purpose of fire drills, for instance, is to enable people to learn how to leave a building in the event of a fire without having to think about what to do.

Although we have not yet developed a computer simulation of this theory, it is based on computational considerations. Such considerations are called "computational" because they are at a particular level of analysis (Marr, 1982), in which knowledge of aspects of the social and physical environment is mapped on to a design for the kinds of operations that could cope with these aspects. Emotions function in real time to redistribute cognitive resources and to manage goal priorities. When an event has been detected that requires re-computing these priorities, an emotion occurs and it helps to manage either the continuation of the current course

of action or the transition to another sequence of action. Emotions help to specify which goals will be actively pursued, and which abandoned, or assigned to a subsidiary or dormant status (see also Stein & Levine, 1990). Emotions have further consequences by way of external signals that co-ordinate group behaviour.

We can summarise the argument so far in three propositions:

Proposition 1. Events and their significance for goals are often unforeseen because: (a) finite organisms cannot be impeccably rational, and they have imperfect models of the world; (b) individuals with several goals are often unable to satisfy all of them simultaneously; and (c) social animals interact together in ways that cannot always be anticipated.

Proposition 2. It follows that junctures in action will occur at which an individual needs to act, but for which there is no fully rational method to select the next action.

Proposition 3. Emotions function to redistribute cognitive resources at junctures in action, particularly where neither cogitation nor reflexes (the residue of fixed action patterns) determine an appropriate course of action. Because some action is probably better than becoming lost in thought, a biologically based system makes ready a small repertoire of actions appropriate to a recognisable type of goal-relevant event. The mechanism tends to constrain the individual to choose the next action sequence from this repertoire. Such a mechanism is a result of natural selection, and the repertoires of actions include both species-specific patterns and individually acquired habits.

WHICH EMOTIONS ARE BASIC?

Many theorists have proposed sets of basic emotions. There are differences among the theories and among the sets of basic emotions that have been proposed. These differences prompt sceptics to argue that it is no longer clear what is meant by the claim that some emotions are basic, and that it has no testable content (see, for example, Ortony & Turner, 1990). Most previous theories, however, have not been based on a functional analysis. Their principal motivation has been to bring order to the disparate set of human emotions by seeking to derive them from a set of basic emotions, e.g. by postulating a set of opposites, by analogy to chemistry or to the mixing of colours (e.g. McDougall, 1926; Plutchik, 1962). Pride, for instance, has been proposed to be a combination of joy and anger; and love a combination of joy and acceptance. Some of the postulates of such theories, however, have no empirical support either subjectively or physiologically, and this again has been noted by sceptics. It is a common

experience to have “mixed” feelings, but this state is characterised by an awareness of alternative and conflicting emotions (see also Ellworth & Smith, 1988; Stein & Levine, 1989). Indeed, our own research (Oatley & Duncan, in press) shows that in more than a third of episodes of happiness, sadness, anger, and fear, a person experiences simultaneously two basic emotions. The most common such mixture is sadness and anger—caused, for instance, by a loss which also frustrates some plan. Our method does not discriminate between true simultaneity and rapid alternation of underlying states. What individuals do not report, however, is the existence of a single emotion made up from phenomenally remote constituents.

According to our theory, emotions are a result of coarse cognitive evaluations that elicit internal and external signals and corresponding suites of action plans. They are emotions because they have cognitive rather than physiological causes. From an analysis of the ontology of simple social mammals, we have proposed the following set of basic emotions: happiness, sadness, anger, fear, disgust (Oatley & Johnson-Laird, 1987), and perhaps desire (Oatley & Johnson-Laird, 1990). Hence, specific emotions are typically caused by the perceptions of general categories of event: (1) happiness with perception of improving progress towards a goal; (2) sadness when a goal is lost; (3) anger when a plan is blocked; (4) fear when a goal conflict or a threat to self-preservation occurs; (5) disgust with a perception of something to reject; and (6) desire with a perception of something to approach. These emotions are indeed basic—however, depending on how the evidence points, other emotions may be basic too. The names of the basic emotions have misleading ethno- and anthropocentric connotations, but in English they come close to suggesting the emotional behaviours of social mammals.

We argue that the status of the basic emotions is corroborated in five ways. First, each of them is an emotion that appears to be universal, and to have universal concomitants, such as a corresponding facial expression (see Ekman, 1973, and Ekman, this issue). Second, each has either a bodily or phenomenological component that can be experienced without the individual knowing the cause of the emotion. Third, the semantics of the large emotional vocabulary of English can be explicated without having to appeal to any other emotions (see Johnson-Laird & Oatley, 1989). Fourth, each term denoting a basic emotion is primitive in the sense that it is semantically unanalysable. It refers to a phenomenological primitive that one needs to have experienced in order to grasp the meaning of the terms. If Mr Spock (of *Star Trek*) does not experience emotions, then it is impossible to explain to him what happiness or sadness are. We could explain what kinds of events are likely to cause these states; we could explain what physiological changes they are likely to bring about and what actions they are likely to elicit. With some perceptual training, such as

experience with Ekman's Facial Action Coding System (Ekman & Friesen, 1978), he would be able to discriminate amongst facial displays of emotions. But we could not explain to him what it was like to *feel* happy or sad, any more than we could explain what red was like to a person who was completely colour blind. Fifth, the apparent complexity of human emotional experience comes from the diverse cognitive evaluations that can elicit and accompany the basic emotions, and that can differ from one culture to another. The accompanying cognitions are also reflected in the vocabulary of emotions. An emotion term accordingly refers to a subset of the basic emotions, typically just a single basic emotion, perhaps with an indication of the intensity of the emotion, as in the series: "contentedness", "happiness", "joy", "ecstasy". A term can also convey that the state has a known cause or object. For example, to use the term, "glad", properly, is to imply a conscious propositional knowledge of what caused the happiness: That one is glad that something has, or has not, happened.

Basic Emotions vs. Components of Emotions

Our theory of basic emotions contrasts with a recent componential proposal made by Ortony and Turner (1990), which is also computationally motivated (see Ortony, Clore, & Collins, 1988). Ortony and Turner reject the hypothesis of basic emotions, and instead they consider it more profitable to analyse emotional expressions and responses in terms of dissociable components that are innate. Their theory is akin to the notion that the underlying components of facial expressions, and other emotional responses, are governed by a system of production rules of the form:

If an Event E1 occurs, then do Action A1
 If an Event E2 occurs, then do Action A2
 . . . and so on.

As an example, Ortony and Turner (p. 332) consider the apparently universal facial expression of anger, which they analyse in terms of separate and dissociable components. We can capture the essence of their claims in the following production rules:

If you become conscious of being unable to attain a goal, then furrow your brow.

If you desire to be aggressive towards the agent responsible for the blockage, then form an open, square mouth that shows your teeth.

If you are determined to remove the source of the goal blockage, then compress your lips.

If you devote considerable attention to the visual environment, then raise your upper eyelids.

These dissociable elements of the prototypical facial expression of anger are invoked by an event in relation to a goal: a goal blockage. This event, however, has attributes such as the “existence of an identifiable agent responsible for the blockage”, which may, or may not, be present in any given episode. Ortony and Turner go on to make a case for dissociable physiological components underlying emotional experiences. “Our view”, they write (p. 322) “is that such differences in physiological responses are usually better interpreted as indicating not so much the presence of specific emotions as the presence of dissociable components of emotions, namely specific appraisals and their corresponding responses”.

We see several problems with a componential analysis of this sort. First, Ortony and Turner allow only “external causes of co-occurrences of sub-components” (p. 323). So, although their system is like a set of production rules, because the independence of sub-components is so fundamental to them, anything that might bind rules together internally is excluded. In contrast to computational production systems which have the power of universal Turing machines, no logical operations between rules are described. Moreover, although alternative environmental events may trigger a single rule, there is no indication that the same event might trigger alternative rules. Hence, there is no indication that their mechanism might generate default operations for states of uncertainty. The system they discuss, therefore, does not address the functional issue of filling the gap between fixed action patterns and impeccable rationality that we have discussed in the previous section. It is hard, indeed, to see how the system differs in principle from the fixed action patterns of insects. Secondly, their account takes a critical step towards treating emotions as a myth of folk psychology. As Ortony and Turner make clear, the dissociable actions that they propose are not caused or linked by anger. On the contrary, the theory dismantles anger into a set of components, which can differ from one case to another. There need be nothing in common to all occasions of anger. How is it possible then for individuals to refer to anger in so many diverse situations? One possible answer is that there is a prototypical set of components underlying all experiences of anger (cf. Fehr & Russell, 1984). Granted a certain number of the characteristic components of anger, then individuals experience the emotion. Such an approach has some plausibility as an account of the concept of an emotion such as anger. However, Ortony and Turner (p. 323) specifically argue against it as an analysis of the emotion itself.

In our view anger typically corresponds to the following sequence:

- an individual's goal is frustrated, often but not necessarily by another agent;
- the individual perceives the blockage;

- a basic anger signal propagates through the cognitive system which the individual experiences as feeling angry;
- as a result of this signal, physiological mechanisms prepare the body for aggression, the face assumes an expression in which the brows are furrowed, etc; and
- plans are made ready for removing the blockage.

The important component in this sequence is the third one: a specific signal propagates which the individual experiences as feeling angry. This component alone is sufficient for an individual to be angry. One can be angry for no known reason, that is, without any awareness of a goal-blockage and without betraying one's feelings by facial expressions or bodily behaviours. If an individual feels angry in such circumstances, then, according to our theory of basic emotions, the feeling is mediated by a primitive unanalysable signal of anger that impinges on consciousness, but without knowing anything else about the state. For a componential analysis, however, a feeling of anger must be mediated by a set of dissociable components that can differ from one such experience to another. Ortony and Turner make this point with great clarity in their discussion of fear (p. 327): "There are various kinds of fear, each consisting of somewhat different components."

It is hard to see what these components of a feeling could be. One putative view is that emotions are valenced experiences, anger is a negative experience, and so the subjective experience is composed of the following components:

Emotion + Negative valence + X

where X is a set of subjective components that distinguish anger from other negative emotions, such as fear. There is a striking dilemma for such a view, however. Either X includes a set of components common to all subjective experiences of anger, or else it does not. If it does contain a common set of components, then we are back once again at a theory of basic emotions: underlying any experience of anger is a common set of components. Hence, on this side of the dilemma, the componential theory is entirely compatible with the theory of basic emotions.

We suspect that Ortony and Turner prefer the other side of the dilemma in which subjective experiences of anger do not contain common elements. In this case, there really is nothing in common to all occasions of anger other than:

Emotion + Negative valence

But these two components fail to distinguish anger from other negative emotions. Hence, this view leads ineluctably to the conclusion that emotions as distinct subjective experiences, such as anger, fear, and sadness,

have no real existence. To reject common components is to reject, not just basic emotions, but all the everyday categories of emotion: One is indeed forced to treat them as a myth of folk psychology.

FOLK THEORIES AND SCIENTIFIC THEORIES OF EMOTIONS

We argued in the previous section that the rejection of basic emotions leads to a rejection of the naïve everyday categories of emotion too. This sceptical view of emotion has always attracted adherents, who regard folk psychology as based on errors that are as egregious as those that underlie naïve physics. Ultimately, according to this form of reductionism, the ideas and terms of folk theory will be replaced by proper scientific explanations. Once again, William James anticipated the critique of folk psychology. Here he is on the pointlessness of studying emotion terms of ordinary language, and of trying to sort them into categories, such as basic and non-basic (James, 1890, p. 485):

If one should seek to name each particular one of them [emotions] of which the human heart is the seat, it is plain that the limit to their number would lie in the introspective vocabulary of the seeker, each race of men having found names for some shade of feeling which other races have left indiscriminated. If we should seek to break the emotions, thus enumerated, into groups, according to their affinities, it is again plain that all sorts of groupings would be possible, according as we chose this character or that as a basis, and that all groupings would be equally real and true.

In more recent times, the argument from this side has gone somewhat as follows: Accounts which include intentional terms, such as “desiring that something”, or “believing that something”, are folk theories that seek to explain and predict individual’s actions. Just as naïve physics depends on the misleading idea of impetus, so folk psychology depends on the misleading idea that beliefs and desires cause behaviour. Newton replaced impetus by coherent laws of motion; so, too, the psychology of belief and desire will be replaced by a proper scientific account of behaviour that will be based, not on such “intentional” concepts, but on the neurophysiology of the nervous system (see, for example, Stich, 1983; Churchland, 1984).

The view implicit in our theory of emotions is that folk psychology is not a myth. It embodies important truths: that individuals have beliefs and desires and needs, that they use their beliefs to decide what to do to attain their goals and then try to carry out these actions—and that emotions have effects on behaviour. An achievement of cognitive science is to rehabilitate mental terms following their banishment during “Behaviourism”, and to show how the psychology of “belief and desire” can be modelled computa-

tionally. There is no warrant for the generalisation from naïve physics to the conclusion that all folk theories are mistaken. In particular, psychological phenomena and physical phenomena are different. A putative account of, say, physical motion is corrigible. But the subjective experience of an emotion is incorrigible in the sense that it is not a hypothesis that could be falsified by evidence in the way that hypotheses about the physical world may be. If you feel definitely happy, you will not be mistaken that you are happy, because, according to us, the feeling of happiness is a direct phenomenological result of a certain kind of signal in the cognitive system. The (folk theoretical) concept which in English is called happiness indicates just such a feeling. It indicates something real. Like sleepiness, or pain, or thirst, it is subjective, not open to consensual validation or evidential refutation.

When an emotion signal does impinge on consciousness it does not have to be interpreted to determine which emotion it represents. It does not represent an emotional state. A conscious emotion is the experience of an emotion signal. Such an experience leaves room for various kinds of doubt, for instance about its cause, about the interpretation of the emotion-eliciting event, about whether the feeling is strong enough to be sure that an emotion really is occurring, or about what kind of emotion it is—particularly if for some reason the emotion is suppressed, or if two emotions occur as a mixture. But, we argue, in straightforward cases where the emotion is felt strongly, e.g. feeling happy at seeing a good friend, feeling angry if someone lets you down, feeling afraid at a traffic accident, there is no doubt about the nature of the emotion itself. So, in the structured diaries of 30 patients attending a gastrointestinal clinic, each asked to record four episodes of emotion of any kind, half of their emotion episodes were experienced in this way. For each episode subjects were asked: “Would you call it a type of any of the following?—happiness/joy, sadness/grief, anger/irritation, fear/anxiety or disgust/hatred.” They were then asked to rate how sure they were about this choice on an 11 point scale from 0 (not at all sure) to 10 (completely sure), (Duncan & Oatley, in prep.). All episodes of emotion were rated as one of the five types. Subjects rated 50.4% of their categorisations as “completely sure”, and only 14% of episodes at 5 or below on the scale of certainty of categorisation. Our claim is also supported by the ready ability of children to learn and to understand the causal sequence of events underlying emotions—the chain from the perception of a goal-related event, to the emotion, and then to a change in action (Stein & Levine, 1989). The cause of the emotion is typically obvious; and this defence of folk psychology is consistent with the existence of basic emotions.

Subjective experiences, of having beliefs, desires, emotions, lie at the heart of folk psychology. As a theory, however, the folk theory of

emotions provides little account of psychological mechanisms, or their physiological bases. The goal of a cognitive science of emotions is thus to spell out a mechanism that is at least consistent with common observations of the causes and consequences of emotions. The persistence over time of these observations does not indicate a stagnation of explanation as a result of isolation from evidence. The evidence is the set of observations of the causes and consequences of emotions to which people are continuously open.

Sceptics might imagine that this hypothesis of a convergence between folk theory and scientific theory is a quirk peculiar to us. But other researchers too, with quite different theories from ours, have come to the same conclusion. Ortony and his colleagues argue that individuals can be usefully consulted about what terms refer to emotions, and that these everyday intuitions map on to the scientific theory of emotions (Ortony et al., 1988). Similarly, Fehr and Russell (1984) and Shaver, Schwartz, Kirson, & O'Connor (1987), have consulted people in a range of ways about their categorisations of emotion terms. These investigators also assume that people know that emotions are caused by certain types of events related to goals. They postulate a correspondence between the results of their studies and scientific categories, and, in the case of Shaver et al., their results support basic categories of emotions, which correspond to some degree to those that we have postulated.

Our hypothesis of a convergence between folk theories of emotions and scientific theories of emotions is, like any other scientific claim, open to refutation. There are indeed several ways to challenge it. One is to argue that self-reports are neither reliable nor valid, and only objective reports of behaviour or physiology should have any part in science. Evidence for such assertions can be derived from the work of Nisbett and Wilson (1977) and Nisbett and Ross (1980), which shows that people are often poor judges of the causes of their judgements and behaviour. The true causes include social conformity, compliance to subtle conditions of experimental designs, and attributional biases. Individuals are not conscious of these factors, and their explanations of their own behaviour ignore them. Instead, they focus on events that are salient, without weighing in any statistically appropriate manner relevant causal factors. They also display pervasive mental shortcomings. They can make gross errors of judgement about the causal effects of their own or others' actions (Jenkins & Ward, 1965), they overlook falsifying evidence (Wason, 1960), and they are biased by information that is more immediately available or that appears to be more representative of the case in hand (Kahneman, Slovic, & Tversky, 1982).

Thus, the argument goes, people do not know the causes of either their behaviour or of their mental states in any way that resembles a scientific account. Not only do they lack a privileged introspective access to how

events cause behaviour, but they are regularly misled by their introspections. They are subject to inbuilt mental deficits in reasoning that will necessarily lead them astray. Hence, folk psychology is not merely irrelevant to scientific theories, but to attend to it is positively misleading.

This kind of argument has encouraged many to eschew evidence based on self-reports, but we believe this is mistaken, for three inter-related reasons.

First, as many of the psychologists studying the shortcomings of the human inferential system have themselves pointed out, their studies deliberately focus on cognitive illusions much as perceptual psychologists seek visual illusions with the goal of revealing the workings of the cognitive system. No psychologist argues from the existence of visual illusions to the claim that all vision is illusory and non-veridical. Likewise, the failures of inference in the psychological laboratory hardly justify the claim that human reasoning is intrinsically irrational (see Johnson-Laird & Byrne, *in press*).

Secondly, as Craik (1943) proposed, the brain models important entities, attributes, and relations in the world. If it had not converged on successful models of important sequences, we would not be able to operate in the world. Thought, behaviour, and communication, are successful more often than not—the central postulates of folk psychology are based on essentially correct, though radically incomplete, mental models. Actions are caused by goals in conjunction with beliefs. The reason, for example, that the engineers in charge at Chernobyl did not report the destruction of the nuclear reactor to the authorities in Moscow is because they did not believe that the reactor had been destroyed. They persisted in the view that the reactor was intact, despite much evidence to the contrary, including the reports of two young probationary engineers whom they had sent to examine it and who paid with their lives for their observations (Medvedev, 1990). Work on inferential failure may reveal causes of such pathological disbelief, but what is clear is that the belief led to a failure to report the scale of the disaster, and that this failure contributed to the appalling delay in evacuating the area.

Thirdly, emotions usually follow immediately after the events that cause them. Therefore, people will not ordinarily suffer the kinds of illusions of thinking just indicated. Such errors occur easily, for instance when causes are probabilistic and temporally distant from effects, as in the studies of Jenkins and Ward. The mechanisms of human learning have been successfully tuned by evolution to sequences in which a causal event is regularly and closely followed by a caused event, as routinely demonstrated in both classical and instrumental conditioning experiments. People are indeed bad in intuitions and judgements made outside this range, but their judgements about emotions derive from many experiences within it. Even

if, as we agree, people do not have introspective access to many kinds of mental process, they can introspect the distinctive phenomenal occurrence of an emotion and they can connect such an occurrence with a putative causal event, which in the typical case is obvious rather than hidden or subtle. Both emotions and their usual causes fall within focal attention. As Ericsson and Simon (1980) argue, it is precisely such data that can be verbalised. Data that are outside attention require inferences of the kind that are subject to the errors pointed out by Nisbett, Wilson, Ross, and others.

LANGUAGE AND THE UNIVERSALITY OF BASIC EMOTIONS

A different argument against the existence of basic emotions concerns language and cross-cultural studies. Wierzbicka (this issue), argues that theorists have assumed the universality of categories and facial expressions that correspond to English terms. This ethnocentricity is immediately revealed, the argument goes, if one takes an emotional term from some other culture and tries to apply it to an English-speaking culture. For example, the Ifaluk emotion of *fago* (translated by Lutz, 1982, as "compassion-love-sadness") seems natural and basic to the culture, but it also seems to have no counterpart in English. Likewise, Lutz describes the Ifaluk emotion called *song*, which she translates roughly as "justified anger". Wierzbicka argues that this emotion does not correspond to any basic notion of anger, and that it should not be referred to by the English word, "anger". This argument is important; and we would like to clarify our position.

When a theorist proposes that the emotion or facial expression of, say, "happiness" is a basic and thus universal emotion, the claim is that among the basic emotions, which have evolved in social mammals and which are experienced and communicated among humans, is one that in English is most closely referred to as "happiness". If we have seemed to imply that the English "happiness" is the basic emotion, we apologise. What we mean is that there is a basic emotion, for which in English "happiness" or perhaps "enjoyment" or, to use Wierzbicka's phrase "something like happy" are the nearest indicators. The underlying emotion can be communicated between people nonverbally, and its communication can be effective despite deep gulfs of language and culture. In another language, the emotional terminology will be different, and whatever term corresponds most closely to "happiness" is likely to differ in its connotations. Thus, on Ifaluk, Lutz describes a concept *ker*, which she translates as "happiness/excitement". Cultural attitudes differ: People on Ifaluk do not believe that they have a Jeffersonian right to the pursuit of *ker*. Although

pleasant, it has a negative social connotation, and people are distrustful of it because it can lead to showing off, and neglect of concern for others which is highly valued on Ifaluk. Nuances of this kind thus reflect different conscious attitudes to each emotion, cultural differences in its causation, and differences in the forms of morally acceptable behaviour to which it may lead. Moreover, most emotion terms in a language have a meaning that combines reference to a basic emotion with other semantic information, such as the cause of the emotion. Thus, for example, "embarrassment" refers to a state corresponding to fear (a basic emotion) caused by finding oneself an object of unwelcome social attention—a common experience in the English-speaking world. Different languages are therefore likely to focus on different causes and objects of emotion, and so emotional terms may be difficult to translate from one language to another. As many philosophers from Quine (1960) onwards have pointed out, when you seek a translation of a word or expression from one language to another, then you must attribute a certain degree of common rationality to the other culture. You are likely to be sceptical about the accuracy of the translation if it implies total irrationality and that you should cease to treat the individuals of the alien culture as having any meaningful beliefs. Indeed, some philosophers go further and argue that complete rationality is a prerequisite if an individual is to be said to hold any meaningful beliefs (Davidson, 1975; Dennett, 1978, p. 20). If emotion terms were fundamentally untranslatable, as Wierzbicka sometimes seems to imply, then it should be impossible for native speakers of incommensurable languages ever to learn one another's terminology. The emotional life of the Ifaluk should remain forever beyond Lutz's empathic grasp. The emotion of the inhabitants of some alien planet may truly be beyond our comprehension, but no such individuals have ever been found on earth. It may be difficult to translate words denoting emotions, but it is not impossible to empathise with a culture and to learn to experience the corresponding emotions.

In short, our general theory of the semantics of emotion terms, which was applied to English terms in the first instance, should be equally applicable to other languages. It preserves the notions—common to both folk theories and scientific theories—that emotions are distinctive states, that they are caused by recognisable events of which people can be consciously aware, and that they can be directed to objects or to other people.

IS THE THEORY TESTABLE?

In this final section, we will counter the criticism that the theory of basic emotions is too vaguely defined to be susceptible of empirical test (Mandler, 1984; Ortony & Turner, 1990). We believe that this criticism is

prompted not by any conceptual difficulty in testing the existence of basic emotions, but by the practical difficulty of such investigations. Indeed, few investigations have been performed that fulfil the conditions to make a compelling case.

One way of falsifying the hypothesis of basic emotions would be to show that the apparent diversity of emotions cannot be reduced to a small basic set because different varieties of, say, fear, have no underlying components in common. What is needed is a set of cumulative studies that test for the universal existence of a small set of basic emotions corresponding to folk theoretical categories. These studies should investigate whether such emotions are experienced, communicated, and recognised universally; and they should investigate whether they have common components in their underlying neurophysiology. Hence, the studies need to examine different cultures, infants on whom culture has yet to impinge, and the physiological systems of animals and human beings (see Panksepp, 1982). It is even possible that certain eliciting conditions for basic emotions will prove to be universal, or at least common to diverse cultures, although the theory does not strictly call for this condition to hold.

Studies of basic emotions are complex, difficult, and time-consuming. Yet, various researchers have begun to undertake them. Ekman and his colleagues have carried out a paradigmatic set of studies that meet the necessary conditions (see Ekman's paper in this issue). They have shown in particular that facial expressions of a basic set of emotions are common across diverse cultural groups, and that basic emotions have distinctive physiological accompaniments.

A stringent hypothesis is that basic emotions should be perceived categorically, just as, for example, the contrast between certain English consonants, has been tested by Etcoff (1990). What distinguishes "bit" from "pit" is a few milliseconds of onset in voicing, i.e. the vibrations of the larynx in the articulation of the phonemes /b/ and /p/. For equal physical differences in voicing onset time, it is difficult to discriminate between two sounds lying on one side or the other of the boundary between /b/ and /p/, but easy to discriminate between two sounds that straddle this boundary. Etcoff argued that if there are basic emotions, then the perception of facial expressions should also be categorical in the same way. Happy faces should be sorted into one category, sad ones into another, and so on. She argued that if she could create equal physical increments in scales ranging between different basic emotions, then there would be categorical boundaries. On one side of them people would see one emotion, on the other a different one, but on either side discrimination should be poorer than across the boundary. She created equal increments using Brennan's (1985) computer program for drawing faces in a way that includes details of eyebrows, eyelids, and mouth. She traced 21 photographs from Ekman and Friesen's

(1976) pictures of facial affect from three models, who each posed expressions of a putative set of basic emotions: sadness, anger, fear, disgust, surprise, and a neutral state. She then used the program to create an incremental series of 11 faces that changed in equal physical increments from one emotion to another. For instance, in one series faces number 1 and number 11 were respectively drawings from the digitised photographs of happy and sad faces of one of the models. Face number 2 derived from the average positions of 10 sets of points from the happy face +1 set from the sad face; face number 3 derived from 9 sets from the happy face +2 sets from the sad face, and so on. With standard psychophysical methods, she then tested the hypothesis of categorical perception of these faces. She observed an abrupt shift in discriminability between the faces in all the series except the one from surprise to fear. She also observed the same effect between the emotion faces and the neutral faces, although the gradations of the neutral faces were more discriminable than those between the emotion faces.

A further corroboration of basic emotions has been obtained by Conway and Bekerian (1987). They found in studies of similarity judgements that emotion terms fell into groups corresponding to basic emotions: happiness/love/joy—misery/grief/sadness—fear/panic/terror—and anger/jealousy/hate. They then used lexical decision tasks to investigate the representation of these concepts in memory. In one experiment, the subjects read two sentences that had previously been judged appropriate to a particular emotion, such as love. They then immediately carried out a lexical decision task in which they were shown a string of letters and had to decide whether or not it was a word. It was either another emotion word from the same basic group, e.g. "joy", an emotion word from another basic group, e.g. "sadness", or a nonword. Interspersed with emotion trials were trials with emotionally neutral filler sentences and words and nonwords. The subjects' reaction times were faster for words from the same basic emotion group than for words from a different emotion group.

These experiments corroborate basic emotions within a single culture. Because the theory postulates an innate and universal foundation for basic emotions, it predicts that the phenomena observed by Etcoff and by Conway and Bekerian should generalise in the same way across different cultures.

CONCLUSION

We have made a case for the psychological reality of emotions and for their foundation on a small set of basic emotions: happiness, sadness, anger, fear, desire, and disgust. Each basic emotion depends on an innate and universal internal mental signal, which can be elicited by rapid and coarse

cognitive evaluations that may be common to diverse cultures. These evaluations concern progress towards goals. The internal signals are causal precursors of subjective experience, somatic change, and plans for action. They are also precursors to external signals, such as facial expressions, that communicate the emotion to others. The theory can be contrasted with the rival hypothesis that there are no basic emotions, but instead more fundamental components, out of which all emotional experiences are constructed (Ortony & Turner, 1990). On the one hand, if there are supposed to be components in common to all subjective experiences of, say, fear, including cases where individuals have no knowledge of the cause of the emotion and react in no outward way to it, then the theory is entirely compatible with basic emotions. On the other hand, if there are not supposed to be any components in common to all subjective experiences of an emotion such as fear, then the theory amounts to a rejection of the folk categories of emotion. Emotions are nothing more than naïve illusions. Once dispelled, they will cease to exist as useful pre-theoretical categories for cognitive science. We have argued that there are no strong grounds for rejecting folk psychology; Ortony and his colleagues have defended a similar position (Ortony et al., 1988). Yet, Ortony and Turner (1990) have questioned both the concept of the basic emotions, and what would count as empirical evidence for or against them. They say that "current uses of the notion do not permit coherent answers to be given to such questions" (p. 329). Their own componential theory, however, seems to be either a variant of the basic emotion hypothesis or else a repudiation of the folk theory. The case for basic emotions has not convinced everybody, but the tests that have been carried out appear to corroborate it.

Manuscript received 4 April 1991

Revised manuscript received 22 November 1991

REFERENCES

- Brennan, S. (1985). The caricature generator. *Leonardo*, 18, 170–178.
- Cherniak, C. (1986). *Minimal rationality*. Cambridge, MA: MIT Press.
- Cohen, L.J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, 4, 317–370.
- Churchland, P.M. (1984). *Matter and consciousness*. Cambridge, MA: MIT Press.
- Conway, M.A. & Bekerian, D.A. (1987). Situational knowledge and emotions. *Cognition and Emotion*, 1, 145–191.
- Craik, K.J.W. (1943). *The nature of explanation*. Cambridge University Press.
- Davidson, D. (1975). *Thought and talk*. In Guttenplan, S. (Ed.), *Mind and language*. Oxford University Press.
- Dennett, D.C. (1978). *Brainstorms*. Cambridge, MA: MIT Press.

- Descartes, R. (1911). *Passions de l'âme*. In E.L. Haldane & G.R. Ross (Eds and Trans.), *The philosophical works of Descartes*. [Originally published 1649.] Cambridge University Press.
- D'Andrade, R. (1987). A folk model of the mind. In D. Holland & N. Quinn (Eds), *Cultural models in language and thought*. Cambridge University Press.
- de Sousa, R. (1987). *The rationality of emotions*. Cambridge, MA: MIT Press.
- Ekman, P. (1973). Cross-cultural studies of facial expression. In P. Ekman (Ed.), *Darwin and facial expression: A century of research in review*. New York: Academic Press.
- Ekman, P. & Friesen, W.V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P. & Friesen, W.V. (1978). *Facial Action Coding System (FACS): A technique of the measurement of facial action*. Palo Alto, CA: Consulting Psychologists Press.
- Ellsworth, P.C. & Smith, C.A. (1988). From appraisal to emotion: Differences among unpleasant feelings. *Motivation and Emotion*, 12, 271–302.
- Ericsson, K.A. & Simon, H.A. (1980). Verbal reports as data. *Psychological Review*, 87, 215–251.
- Etcoff, N. (1990). *Categorical perception of facial expressions*. Paper presented to the Fifth Annual Meeting of the International Society for Research on Emotions, Rutgers University NJ, 25–28 July.
- Fehr, B. & Russell, J.A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General*, 113, 464–486.
- Frijda, N.H. (1987). Comment on Oatley and Johnson-Laird's 'Towards a cognitive theory of emotions'. *Cognition and Emotion*, 1, 51–59.
- Hammond, M. (1990). Affective maximization: A new macro-theory in the sociology of emotion. In T.D. Kemper (Ed.), *Research agendas in the sociology of emotions*. Albany, NY: State University of New York Press.
- Henle, M. (1978). Foreword to R. Revlin & R.E. Mayer (Eds), *Human reasoning*. Washington: Winston.
- James, W. (1890). *The principles of psychology*. New York: Holt.
- Jenkins, H. & Ward, W. (1965). Judgements of contingency between responses and outcomes. *Psychological Monographs*, 79, No. 594.
- Johnson-Laird, P.N. (1983). *Mental models*. Cambridge, MA: Harvard University Press/Cambridge University Press.
- Johnson-Laird, P.N. & Byrne, R.M.J. (1991). *Deduction*. Hove: Lawrence Erlbaum Associates Ltd.
- Johnson-Laird, P.N. & Byrne, R.M.J. (In press). Models and deductive rationality. In K. Manktelow & D. Over (Eds), *Rationality*. London: Routledge.
- Johnson-Laird, P.N. & Oatley, K. (1989). The meaning of emotions: Analysis of a semantic field. *Cognition and Emotion*, 3, 81–123.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under uncertainty: Heuristics and biases*. Cambridge University Press.
- Lutz, C. (1982). The domain of emotion words on Ifaluk. *American Ethnologist*, 9, 113–128.
- Mandler, G. (1984). *Mind and body: Psychology and emotions and stress*. New York: Norton.
- Marr, D. (1982). *Vision*. San Francisco, CA: Freeman.
- McDougall, W. (1926). *An Introduction to social psychology*. Boston: Luce.
- Medvedev, Z.A. (1990). *The legacy of Chernobyl*. New York: Norton.
- Nisbett, R.E. & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgement*. Englewood Cliffs, NJ: Prentice Hall.
- Nisbett, R.E. & Wilson, T.D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259.

- Oatley, K. & Duncan, E. (In press). Structured diaries for emotions in daily life. In K. Strongman (Ed.), *International review of studies of emotion*, Vol. 2. Chichester: Wiley.
- Oatley, K. & Johnson-Laird, P.N. (1987). Towards a cognitive theory of emotions. *Cognition and Emotion*, 1, 29–50.
- Oatley, K. & Johnson-Laird, P.N. (1990). Semantic primitives for emotions. *Cognition and Emotion*, 4, 129–143.
- Ortony, A. & Clore, G.L. (1989). Emotions, moods, and conscious awareness: Comment on Johnson-Laird & Oatley's "The language of emotions: An analysis of a semantic field". *Cognition and Emotion*, 3, 125–137.
- Ortony, A., Clore, G.L., & Collins, A. (1988). *The cognitive structure of emotions*. Cambridge University Press.
- Ortony, A. & Turner, T.J. (1990). What's basic about basic emotions? *Psychological Review*, 97, 313–331.
- Panksepp, J. (1982). Towards a general psychobiological theory of emotions. *Behavioral and Brain Sciences*, 5, 407–467.
- Plutchik, R. (1962). *The emotions: Facts, theories, and a new model*. New York: Random House.
- Quine, W.V.O. (1960). *Word and object*. Cambridge, MA: MIT Press.
- Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology*, 52, 1061–1086.
- Simon, H.A. (1967). Motivational and emotional controls of cognition. *Psychological Review*, 74, 29–39.
- Stein, N.L. & Levine, L.J. (1989). The causal organisation of emotional knowledge: A developmental study. *Cognition and Emotion*, 3, 343–378.
- Stein, N.L. & Levine, L.J. (1990). Making sense out of emotion: The representation and use of goal-structured knowledge. In N.L. Stein, T. Trabasso, & B. Leventhal (Eds), *Psychological and biological approaches to emotion*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Stich, S. (1983). *From folk psychology to cognitive science*. Cambridge, MA: MIT Press.
- Stich, S. (1990). Rationality. In D.N. Osherson & E.E. Smith (Eds), *Thinking: An invitation to cognitive science*, Vol. 3. Cambridge, MA: MIT Press.
- von Uexküll, J. (1957). A stroll through the worlds of animals and men. In C.H. Schiller (Ed. and Trans.), *Instinctive behavior: The development of the modern concept*. London: Methuen.
- Wason, P. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, 12, 129–140.