

NATURAL EXPRESSION : ITS PLACE IN
COMMUNICATION AND ACTION

K.G. Rhys-Jones

Ph.D. Thesis
Bedford College

ProQuest Number: 10098471

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10098471

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

For Ellie

ABSTRACT

The dominant aim of this thesis is to define a concept of natural expression the precise extension of which is left undetermined, being subject to empirical investigation. Nevertheless, paradigm examples with respect to which the question is raised : 'Are they natural expressions?' include smiling and other facial expressions.

In pursuit of this aim suggestions are made concerning the logical form of sentences about expression (expression is not, in my view, best regarded as a relation), comparisons are drawn between the interpretation of utterances and the interpretation of non-linguistic expressions (and both contrasted to explanation), and, subsequently, between the latter and what Davidson calls the 'causal redescription of actions'.

The task of defining expression is then conceived of as partly consisting in the specification of the cognitive effects such that when an action is truly described 'in terms of' them, it is expressive. In the ensuing analysis, much of the work is done by adapting Christopher Peacocke's notion of 'prima-facie evidence'.

The notion of 'physiognomic perception', ie. the way in which we see eg. anger in eg. a scowl, is critically discussed and related to the foregoing.

The relevance of correlations between emotion and behaviour to natural expression is then assessed.

In the penultimate chapter I argue that expressive behaviour is natural if there exist innate dispositions to exhibit it under certain conditions and innate capacities to interpret it correctly. The relevance of this to radical interpretation, concept-acquisition, and the explanation of action is considered. The contributions of some social scientists to this area are critically examined and broad prescriptions are made concerning the nature of the evidence to be taken into account in determining the naturalness of a given type of behaviour.

I conclude with a tentative endorsement and formulation of 'immanent realism' with respect to other minds.

CONTENTS

ACKNOWLEDGEMENTS	6
I IS EXPRESSION A RELATION?	7
1. Introduction	7
2. Expression as a relation	9
3. Alternative relata	16
4. Expressive properties	17
5. Expressing that ...	22
II INTERPRETATION AND EXPLANATION	28
1. Utterances	28
2. Facial expression and the linguistic analogy	38
3. Describing behaviour	40
4. Describing perceptual experiences	44
III THE EFFICACY OF EXPRESSIVE BEHAVIOUR	48
1. Behaviour and bodily movements	48
2. How basic are interpreting descriptions?	54
3. Causation and expression	59
4. Conclusion	65
IV COMMUNICATION	66
1. Expression as potential communication	66
2. Gricean communication	70
3. Communication and belief	77
4. Communication and evidence	82
5. Some objections	88

V	PHYSIOGNOMIC PERCEPTION	92
	1. Physiognomic perception and interpretation	92
	2. Inference versus perception	95
	3. 'Seeing-in'	99
	4. Physiognomic versus representational seeing	107
	5. Seeing-as and physiognomic illusion	112
VI	CORRELATION AND APPROPRIATENESS	121
	1. Performance and correlation	121
	2. Performance and appropriateness	125
	3. Correlation as a necessary condition	128
	4. Interpretation and correlation	137
	5. Interpretation and appropriateness	141
VII	NATURAL EXPRESSION	145
	1. Justification of the subject-matter	145
	2. Is any expressive behaviour universal?	152
	3. Inter- and intra-cultural interpretation	159
	4. The innateness of physiognomy	166
VIII	CONCLUDING NOTE ON THE PHILOSOPHICAL RELEVANCE OF NATURAL EXPRESSION	175
	APPENDIX : SPONTANEITY AND PURPOSELESSNESS	181
	BIBLIOGRAPHY	188

ACKNOWLEDGEMENTS

I am indebted to David Wiggins, who first suggested the topic for this thesis, and to Brian O'Shaughnessy for helpful discussion at an early stage of its inception. The actual writing was carried out under the supervision of Mark Sainsbury. I am very grateful to him for reading and commenting, both philosophically and stylistically, on two draughts of this essay, for his advice in general and for his tactful encouragement.

I am also most grateful to my parents and to Eleanor Edwards for material and other kinds of support at various times over the last three years and to the latter, again, for her rapid and excellent typing of the final manuscript.

I. IS EXPRESSION A RELATION?

1. Introduction

An initial survey of the diversity of ways in which the term 'expression' and its cognates, 'expressive' and 'to express', are used is likely to induce considerable perplexity in the aspirant philosophical analyst and certainly ought to induce despair in anyone who hopes to isolate and elucidate *the* concept of expression. There is not one, but there are several such concepts, some of which are more deserving of philosophical attention than others. The aim of this thesis is to elucidate in some detail one concept of expression - a concept which is operative when we say that a smile is an expression of joy, or that a scowl is an expression of anger. A precedent for this choice of paradigm is to be found in Wollheim's articles of 1964 and 1966. The primary criterion of adequacy which the following discussion is intended to meet is that it should be true at least of a certain range of episodic facial expressions, which includes smiles and scowls. Further applications of the ensuing analysis are to be welcomed as a bonus, but are not absolutely required for its confirmation. Nevertheless, there is room for reasonable optimism about its applicability to some gestures such as fist-shaking, foot-stamping, etc., and, with relatively slight modifications, to gait and posture and physiognomic 'sets' which express enduring psychological traits or propensities like cruelty and benevolence. This may appear to be rather an odd restriction to place on the subject-matter of what, after all, purports to be a philosophical thesis. Insofar as a justification really is called for, it lies, at least in part, in the ontogenetic significance of facial expression in all human communication.¹ Secondly - and it is very plausible that the two points are connected - if, as Wittgenstein ((1958)§54) thought, there are characteristic signs of correcting a slip of the tongue, the recognition of which is not impeded by ignorance of the speaker's language, or, as Quine thought ((1960)Chapter 2), we can assume the ethnologist's ability to recognise the behaviour of native speakers of an alien language as indicative of

1. cp Taylor (1979) p.78. 'The smile plays a crucial role ontogenetically in our being able to enter into communication as human beings in the first place.'

assent or dissent, then these types of behaviour belong to an important general category which also includes smiling.

The conspicuously unsystematic proliferation of concepts grouped under the heading of 'expression' apparently exceeds any plausible or interesting explanation of their generation even by "family resemblances". Recognition of this requires the abandonment both of any reformist attitude towards the nebulously defined category of "ordinary language" and of any aspiration towards completeness. It is pointless to complain, but salutary to observe that, in Goodman's phrase, we ordinarily 'play fast and loose' with the word 'expression'. A protest, therefore, is to be recorded against those who, like Collingwood ((1938) pp. 121-124), urge a distinction between "betraying" and "expressing" emotions which requires us to refrain from calling laughter an expression of mirth and against those who, like Danto ((1973) Chapter 6), insist on the contrary that it is only such "spontaneous" manifestations of the inner life that should be classified as expressive. An exhaustive description of the meaning of 'expression' would take the form of an extensive inventory, interesting mainly to lexicologists and not to philosophers.² Accordingly, when I have occasion to deny that a kind of event or object is an expression of some other kind of (mental) event, although they may be otherwise related, this should be taken to mean that the event or object in question is not to be described as expressive in the sense in which - and only in that sense - smiling and scowling, or smiles and scowls are properly describable as expressive.

Despite the chaotic diversity emphasised in the preceding paragraphs some philosophers have, with a measure of success shortly to be estimated, distinguished two major categories of expression which correspond approximately, but only approximately, to the difference between artistic and behavioural expression. However, the kind of bipolar classification put forward by Goodman ((1976) p. 45f) and Tormay ((1971) especially Chapter 2) is put under considerable pressure, at the level of logical form to which it applies, by precisely the kind of example which interests me. The surface grammar of sentences about expression is relational. It encourages construal in terms of

2. A recognition of this diversity is to be found in Sircello (1972), but it has, in my opinion, led Sircello to an over-fastidious attention to nuance.

an expressive and an expressed object; an 'expressans' and an 'expressandum' (Aldrich (1978) p. 203). In the case of behavioural expression, the relational form of the surface grammar is retained as a reliable guide to its logical form. I will argue in Section 2 of this chapter that while this approach seems to work for some cases, it tends to justify negligence with respect to the expressive properties of some behaviour and cannot easily accommodate insincere expression. In Section 3, I briefly investigate some ways of retaining the relational form of expression, but modifying the ontology of its relata. In Section 4, I examine proposals to eliminate the expressed object initially in their application to artistic expression. I then consider the extension of this procedure to the recalcitrant cases of behavioural expression and argue that it fails because of a highly implausible assimilation of insincere to fictitious or other kinds of "non-inferential" expression (cf. Tormey (1971) Chapter 2). In the final Section, I introduce an altogether different conception of the logic of expression to deal with smiling, scowling, etc.. The basic proposal is to treat 'expresses' as a kind of - or as based on - a sentence operator.

2. Expression as a relation

2.1 Sentences which take the form 'x's ϕ -ing expressed F-ness'³ appear to report a two-place relation between an action and an emotion.⁴ Given a general account of the logical form of action sentences something like the one in Davidson (1967), sentences of the form 'x expressed F-ness by ϕ -ing' appear to report the holding of a three-place relation between an agent, an action and an emotion. At this stage, it will be sufficient to concentrate on the two-place relation alone. There are various strategies that could be adopted in order to interpret sentences of the kind in question. I will start by exploring the consequences of taking seriously the idea that expression is a

3. When the subject taken by 'expresses' is an action, object or non-intentional event, I will regard it as interchangeable with '... is expressive of ...' and '... is an expression of ...'.

4. In general, I will be committed to the existence of mental items. I will talk of mental or affective events and states and assume that emotions, feelings, etc. are, like desires and beliefs, examples of them. I do not discriminate sharply between events and states. What governs my choice on any particular occasion will be no more than an intuitive feel for the duration of the item in question, states being less transitory than events.

relation and treating possible replacements of 'F-ness' as names of emotions.⁵ These consequences are exhibited particularly clearly in Tormey (1971), but they are also present in many traditional accounts of artistic expression as a relation (usually causal) between a work of art or certain properties of it and the artist's feelings. (This point is discussed in Wollheim (1968) section 15f.) If 'x's ϕ -ing is an expression of F-ness' describes a relation between an action and an emotion, then it cannot be true unless the emotion exists. An emotion exists if and only if there is someone who has or feels or suffers from that emotion. Sadness, for example, exists when someone has a feeling of sadness and not otherwise. And people have feelings of sadness if and only if they are sad. So a person cannot express sadness unless he or she is sad; 'x's ϕ -ing is an expression of F-ness' strictly entails 'Fx'. Thus the conception of expression as a relation which holds between token actions and token mental events is all that is needed to establish Tormey's 'inferential' concept of expression (Tormey(1971) p. 59f). The task for a theory of expression thus becomes the elucidation of this relation, involving decisions, for example, as to whether or not it is causal and a description of the causal route from emotion to behaviour.

It is important to stress that if expression is a relation between an emotion felt by a person and an action of which that person is an agent, or an unintended movement of the person's body, or an object or event made or produced by the person, then nothing about the nature of this relation, as distinct from other relations which might hold between similar items, follows from the observation that the ascription of the expressive action or object to a person entails the attribution of the expressed emotion to that person. This point needs to be mentioned because, though seemingly elementary, it is often over-looked in the intricacies of philosophical debate about the relation between mental events and behaviour. Tormey's apparent implication that the non-contingency of statements about the relation between an expressans and an expressandum indicates some special feature of the relation itself is comparable to a certain kind of argument which is designed to show that mental events such as desires cannot be the causes of behaviour (eg. Taylor (1964) p. 33,

5. Replacements for 'F-ness' will not, of course, always include the suffix '-ness'. Its function in the schematic context is to indicate the relation which 'happiness' bears to 'happy' which is, semantically, the same as the relation born by 'anger' to 'angry'.

Kenny (1963) p. 109). The point can be seen to turn on the opacity introduced by modal operators. 'It is necessary that if a ϕ -ing is performed then its agent is F' plainly does not sanction an inference, even given identity of the ϕ -ing with a ψ -ing, to : 'It is necessary that if a ψ -ing is performed then its agent is F'. If this were a reasonable inference, then any causal relation between two events such that the first could be picked out by a description of it as the cause of the second (and this is surely true of all causal relations) would not be a causal relation. This is plainly contradictory. One could not argue, therefore, that because 'if x ϕ 's and ϕ -ing is an expression of F-ness then Fx ' is necessary, then there is a certain kind of relation, eg. causation, which cannot hold between instances of ϕ -ing and instances of F-ness. At its simplest, the point is that one should avoid confusing relations between sentences with relations between things. One of the causes of this kind of muddle is allowing one's variables to range indiscriminately over descriptions and their descripta. If one refrains from doing this, then the question : 'Is the relation between X and Y causal or logical?' is readily seen to be nonsensical. Therefore the question as to whether 'x's ϕ -ing is an expression of F-ness' entails 'Fx' is equivalent (subject to a qualification to be introduced shortly) to the question of whether it expresses a relation.

2.2 An example which appears to meet the requirements of Tormey's strictly "inferential" concept is the following : 'John expressed his sympathy by smiling', or (i) 'John's smiling expressed his sympathy'. Tormey correctly maintains that 'It would seem to be, prima-facie, meaningless to say that I am expressing my belief when there is admittedly no such belief that can be attributed to me' (Tormey, *ibid*, p. 68n). Analogously, it would be contradictory to assert that John expressed his sympathy and to deny that there was any sympathy felt by John. But this fails to establish that 'expression' describes a relation between the items mentioned, ie. John's smiling and John's sympathy. The failure is due to the entailment being adequately accounted for by the presence of the possessive pronoun, 'his'. The verb, 'expressed', in 'John expressed his sympathy', could be replaced by any other verb - eg. 'concealed', 'talked about', 'ignored' - without removing the implication that John felt sympathetic. There is an analogy with reference to objects literally possessed or owned by the subject. 'John sailed his boat' is true only if there is a boat which belongs to John. The same applies to : 'John dreamed about

his boat', or : 'John dreamed that he was sailing his boat'. But there is no temptation to infer from this that there are two senses of 'dreams', one of which justifies inferring the existence of *y* from the fact that *x* dreams about *y*, and one of which does not - or such that one of them allows quantification into the context introduced by 'dreams that ...' and the other does not. The explanation of the implication lies in the fact that sentences of the form '*x* ϕ 's *x*'s *y*' are equivalent to conjunctions, one of whose conjuncts asserts the possession of *y* by *x*.⁶ No doubt the idiom of possession is applied metaphorically to people and their emotions. Its logical pattern, however, is unaltered. It is simply carried over in its entirety from the domain of owners and the commodities they own to the domain of persons and the beliefs, desires, emotions, etc. which are attributable to them. However, considerable pressure is generated here to consider expression as an existentially quantifiable relation. The same pressure is not generated by sentences which lack a possessive pronoun. This strongly suggests that there *are* two senses of 'expression'. A similar problem arises for many propositional attitude sentences, but we are prepared to tolerate a high degree of complexity in their analysis because of the strength of our initial intuition that there is only one concept of belief or of dreaming (cp. Quine (1956) p. 188 in (1976)). A parallel intuition with respect to expression appears not to be very widespread. To avoid complicating matters unnecessarily, it may, therefore, be better to rest with a relational concept of expression applicable where the emotion expressed is actually felt and devise a totally autonomous concept for other cases.

This approach will be reinforced if examples are available which fit the relational pattern without depending on the antecedently established existence of the emotion expressed. An example which could plausibly be interpreted as meeting this requirement is the following :

(ii) William expressed resentment by staying away for three days.

(ii) appears to differ from a third kind of example in the way in which one can imagine it being used.

(iii) John expressed sympathy by smiling.

6. The possessive pronoun does not always function in quite this way. A different interpretation of 'his' is given in Section 5, but I see no alternative in the present context.

(i) and (ii) seem to exemplify Tormey's inferential concept of expression while (iii) does not, or at least need not. It is not obviously contradictory to deny that John felt sympathetic while assenting to (iii). But do (i) and (ii) differ in the same way from (iii)? I submit that they do not, and that there is some feature shared by (i) and (iii), but not by (ii). There is some intuitive support for this. If it is true that William expressed his resentment by staying away for three days, it may also be true that, in a different sense of 'expresses', William stayed away not only because he felt resentful, but also because he was too shy or too polite to express his resentment. In the first sense of 'expression', expression₍₁₎, William's staying away was an expression of resentment. In the second sense, expression₍₂₎, resentment can only be expressed by such activities as scowling, sulking, rude utterances and so on. It appears that there is something that John does in both (i) and (iii) which is the same, and which can be described as expressing sympathy, and which is not equivalent to what William is reported to have done in (ii) with respect to resentment.

Expression₍₁₎ describes the relation between William's resentment and William's staying away for three days. If one admits causal relations between mental states such as resentment and behaviour such as staying away, then it seems that expression₍₁₎ is reducible wholesale to this kind of causation. This seems to conform to Tormey's inferential concept of expression which he is prepared to apply to the relations between desires and the actions to which they give rise and to the relations between actions and the beliefs which justify them (ibid. p. 30). This paves the way for the claim that there is no 'descriptively distinct class of behaviour that constitutes expression' (ibid. p. 44). But there is no very pressing need for such a concept, since it does not enable us to say anything that could not equally well be said in terms of causation or explanation. The important point here is that expression₍₁₎ is not only inappropriate to (iii), but also to (i). For while (i), unlike (iii), asserts some kind of (eg. causal) relation between John's smiling and his sympathy, it is not thus fully characterised. What is missing, unsurprisingly, is that (i) also entails (iii). There is no equivalent entailment by (ii). An approximate account of the differences between the three examples mentioned can be given in terms of the difference between tokens and types. Thus (i) asserts the conjunction of a relation between an action-token, John's smiling, and an emotion-type, sympathy,

with a relation between John's smiling and an emotion-token, John's sympathy. It presupposes an association of a certain type of action, smiling (sympathetically), and a type of emotion. (ii) asserts only a relation between an action-token, staying away, and an emotion-token, resentment. It presupposes no antecedent relation between the types of which the items mentioned are tokens. (iii) asserts only a relation between an action-token and a type of emotion, implying nothing as to whether the latter is tokened.

(i), therefore, appears to combine elements taken from (ii) with elements taken from (iii). It seems that the analysis of the first must be postponed until after analysis of the third. The latter will re-import the expressive properties of behaviour like smiling sympathetically which are effectively eliminated by the construal of (i) on the model of (ii). The problem faced by the token-relational account is initially exposed by its failure to accommodate (iii). In other words, it cannot account for expressive behaviour in a way that leaves it open whether or not the behaviour is sincere. It works only for examples like (ii), where the authenticity of the emotion is assured. It makes sense to ask whether a smile is sincere. It does not make sense to question the sincerity of staying away for three days. Tormey is prepared to equate the questions : 'Is x's ϕ -ing an expression of F-ness?' and : 'Is x's ϕ -ing a sincere expression of F-ness?' in the domain of behavioural, including linguistic, expression (ibid. p. 68). What is needed here, however, is a concept which will generate an intelligible interpretation of the following, obviously reasonable remark of Anthony Kenny's :

An assertion ... may be the expression of a belief. Of course, an assertion may be a lie. In that case it expresses a belief which the utterer does not have. (Kenny (1976) p. 40)

And, of course, a smile may be insincere, in which case, it expresses a feeling the smiler does not have. Moreover, because (i) entails (iii), the relational concept fails to account for sincere as well as insincere expression. In short, it cannot explain expressive behaviour of the kind that is sincere or insincere, but only behaviour which cannot be evaluated on this scale. A supplement to the theory is required, and the most natural source in which to search for this supplementary material is the theory of artistic expression, which provides the only other concept of expression available to those who opt for a bi-polar general theory.

2.3 Before proceeding to an examination of the prospects for a

relational theory thus supplemented by a theory of expressive properties, a possible objection to the arguments of the preceding paragraphs needs to be recorded. The argument depends on the assumption, which is, prima-facie, plausible enough, that where 'F-ness' ranges over emotions such as anger and 'F' over properties such as the property of being angry, 'x has a feeling of F-ness' is equivalent to 'x is F'. Given this assumption, to show that on a relational construal 'x expresses F-ness' entails that x has F-ness is simultaneously to show that it entails that x is F, and hence that the relational concept corresponds to relatively few of the interesting things that can be said using the term 'expresses'. (As always, 'x' ranges over people and not their performances or what they achieve by them.) This assumption has been questioned by Wollheim ((1968) p. 45 in the 1970 edition). Wollheim maintains that there are ways in which a man can stand in a 'possessive relation' to his feelings, eg. fear, which do not entail that he is actually afraid. We can experience feelings in something other than the 'primary sense'. Recognition of this is certainly required if an adequate phenomenology of emotion, and particularly of the role of emotion in both the spectator's and the creator's relation to the work of art, is to be achieved. The primary and secondary senses in which a man can be said to have a feeling correspond, perhaps, to the difference between having a belief and merely entertaining a thought. Thus, the objection continues, a man can express the belief that p even if he does not believe that p, so long as he stands in a secondary possessive relation to that belief. His utterance then stands in an expressive relation to the belief which he possesses, albeit only in a secondary sense (the thought which he entertains).

There are two reasons why I think this objection is weak. First, the suggestion upon which it is based is, as it stands, very obscure indeed. It is merely adumbrated by Wollheim and there is no reason to suppose that, given the metaphorical sense of 'possession' operating here, a detailed theory of this aspect of emotional experience would support the relational theory of expression. Anyway, the onus is squarely on the defender of that theory to expand the suggestion before any use can be made of it. Secondly, I doubt whether Wollheim's observation is relevant in the present context. Though prima-facie plausible in the aesthetic context, it is highly unlikely in the behavioural context that a person stands in any relation to an existing feeling when his behaviour is insincere. The aesthetic context

includes the stage and it is often alleged that a good actor is obliged to feel the emotions of the characters he plays. The same hypothesis is rarely put forward to explain mundane hypocrisy about which Aldous Huxley once remarked that it is merely bad acting.

3. Alternative relata

A method of avoiding the less desirable consequences of the relation theory could be based on an ontological modification of what is expressed. Words like 'sadness', 'anger', etc., exhibit a certain ambiguity - an ambiguity not shared by words like 'redness', 'largeness', etc.. Both are substantives which are related systematically to predicates in such a way that the predicates are in some sense more semantically primitive. But terms like 'sadness' and 'anger' acquire a life of their own as names of states, or occurrences, which people can be in, or which happen in or to people. These terms function as names of emotions at least some of the time. But there is no equivalent kind of individual which is ever named by a substantivised colour adjective. However, substantive emotion words also, some of the time, function in the same way as substantive colour words. Sometimes, 'sadness' is the name of a sad thing - an emotion. At other times, 'sadness' stands to 'sad' as 'redness' stands to 'red'. In order to preserve the relational concept of expression, all that is required is a broadly nominalist strategy for the interpretation of substantivised adjectives.

If a neat solution to the problems of Section 2 could be achieved by platonising, then perhaps parsimony should not be insisted upon. The proposal would be to postulate the existence of universals to replace individual emotions in cases of expressive action, where no emotion is genuinely felt by the agent. However, there are grounds for pessimism about the prospects for this approach. The goal is to construct an integrated theory applicable at least to (i) and (iii) discussed above, if not to (ii) also. Expression is to be construed in (iii) as a relation holding between a piece of behaviour, John's smiling, and a universal, sympathy. (ii), accordingly, is to be construed as a relation of the same kind holding between a piece of behaviour, William's staying away for three days, and an emotion, resentment, actually suffered by William. But, while the notion of the relation (eg. identity) holding between a pair of universals being of the same kind as the relation holding between a pair of individuals is at least comparatively perspicuous, the notion of the same relation

holding between an individual and either another individual or a universal is thoroughly obscure. Still, the requirement that the same relation is exemplified in both (ii) and (iii) can be dropped, since it is not unreasonable to assume that if there is any concept of expression applicable to (ii), it is not the same concept as that which operates in (iii). This still leaves (i), where John expresses the sympathy he actually feels by smiling. It looks as though (i) instantiates both kinds of relation at once. This is likely to prove cumbersome, to say the least, since it raises the further question of how the two relations are related. I will not pursue this. The same difficulties attach to the identification of the referent of 'sympathy' in (iii) as a type or, with added implausibility, as a class.

There is, however, no absolute injunction against naming properties as such, insofar as this can be achieved by higher order quantification. And quantification over properties is not straightforwardly reducible to ordinary quantification; witness the non-equivalence of ' $\exists F (\phi Fx)$ ' and ' $\exists x (\phi Fx)$ '. This might seem to encourage an attempt to construe expression as a relation between behaviour and emotional properties as such. However, such a proposal both seems to be open to the same objection as the previous one to treat expression as a relation between abstract objects and behaviour and is faced with the further problem of showing that properties (emotional or otherwise) can exist independently of there being any individuals which bear them. If it fails in the latter task, then the objectionable entailment from 'x's ϕ -ing expresses F-ness' to 'Fx' will not have been effectively blocked.

4. Expressive properties

4.1 In the philosophy of art, there is a school of thought according to which, a work of art is expressive not because it is somehow related to the artist's state of mind anterior to, or during its production, but in virtue of certain features it possesses and which are accessible to viewers, however ignorant of the artist's biography. A well-known statement of the position is the following by Bouwsma :

As the life and the light describe (respectively) the geranium and the sun, so too does the sadness describe the music. (Bouwsma (1971) p. 248-9 in Hospers (1971))⁷

The point is perhaps rather oddly put, but it is reasonably clear and it has an obvious application to facial expression. When a man smiles, his face is happy. When he scowls, his face is angry. The man whose face it is may be neither happy nor angry. (Wollheim (1966) pp. 90 and 99 in (1973), Goodman (1976) p. 50.) The predicates '... is happy', '... is angry', clearly function here in a way which is not equivalent to the way in which they function when they are satisfied by feelings or by people who have those feelings. A face, or a work of art, cannot feel happy or angry, but they can, in a sense, be happy or angry. Predicates used in this way have been designated in various ways by philosophers who have noticed them; 'physiognomic' (Hampshire (1960) p. 153 in (1972) and Wollheim (op. cit.) p. 92); 'anthropomorphic' (Sircello (1972)); and, wrongly I think, 'metaphorical' (Goodman (op. cit. p. 51). By extension, certain verbs, eg. 'to scowl' and certain descriptions of facial configurations, eg. 'a scowl', are physiognomic in that, although they do not explicitly mention anger, they entail descriptions which do.⁸ "'To scowl" is to "produce an angry lie of the face".' (Wollheim (op. cit.) p. 97) There are no scowls which are not angry lies of the face, though there may be angry lies of the face which are not scowls. 'Scowls', then, is an implicit physiognomic predicate. 'Angry', applied to a face, is an explicit physiognomic predicate.

The physiognomic role of some predicates must be recognised, but it is far from certain that a physiognomic description can be found for every case of artistic or, more generally, non-inferential expression, nor that any important clarification could be achieved by doing so. A work of art might express a profound longing for the Kingdom of God. Does this make it a profoundly-longing-for-the-Kingdom-of-God work of art? In certain circumstances, a face might express an

7. See also Gurney (1971) p. 117 in Hospers (op. cit.) and Beardsley (1971) p. 302, also in Hospers (op. cit.). A discussion of this view is to be found in Aldrich (1978) pp. 203f.

8. I am not sure that 'smiles' is a physiognomic predicate in this sense. Although 'x smiled' apparently implies that x's face was happy, I think that this implication is conversational in Grice's (1961) sense. One can also smile sadly. (I am indebted to Mark Sainsbury for insisting on this point.)

avid desire for caviar. Does this make it an avidly-caviar-desiring face? Even if the invention of complex predicates like these is tolerated, it is not clear that they serve any purpose in the absence of an analysis of the notion of a physiognomic description. For Bouwsma is surely wrong to claim that '... is sad' is satisfied by a piece of music in exactly the same way as '... is living' is satisfied by a geranium. In fact, the concept of a physiognomic predicate can be elucidated only by appealing to the concept of expression it was introduced to explain or, in the case of Beardsley (op. cit.), to replace. A lie of the face or a piece of music is sad if and only if it expresses sadness. A predicate, F, when it occurs physiognomically, means 'is expressive of F'.

Goodman has made an attempt to characterise physiognomic predicates without mentioning expression. According to Goodman, a work of art expresses F-ness only if 'F' metaphorically denotes the work of art. (Goodman (1976) pp. 50-52. At this stage, I am concerned only with the first half of Goodman's analysis.) A piece of music, a face, or a gesture⁹ expresses sadness or joy only if it is metaphorically sad or joyful. But the idea that sadness is even metaphorically ascribed to faces or works of art faces considerable intuitive resistance. When emotions are metaphorically attributed, they are attributed to entities, such as machines, which we do not believe to be capable of genuine mental activity. To say, in the case of machines, that emotion-predicates denote metaphorically, is to say that mental states are metaphorically ascribed to them. But mental states are never in any sense, metaphorical or otherwise, attributed to faces or works of art, though they are sometimes appealed to in their explanation. To say that an item of this kind metaphorically feels sad is only minimally less absurd than saying that it literally feels sad. Of course, in insisting that physiognomic predicates are literal, I do not mean to deny that they are somehow derivative. To that extent I agree with Goodman. But this derivation cannot be explained without using the operative concept of expression.

The 'descriptivist' programme, therefore, is best understood as an attempt to explain artistic expression; not as an attempt to abolish it. Descriptivism is designed to meet the need for a

9. On p. 58, Goodman indicates that he takes his theory to be applicable to facial and gestural as well as artistic expression.

non-relational concept of expression applicable primarily to art. But this need is not met by focusing on a straightforwardly one-place predicate, 'expressive' (as Tormey seems to do at, eg. p. 109). The aim is to block the question: 'What does the work of art express?' only insofar as this question invites a token-relational interpretation. But it can also be construed, as Aldrich puts it, as a request for further specification of the expressive property or properties attributed (Aldrich (op. cit.) p. 204). Any description of a work of art as expressive presupposes the availability of such a specification, or would presuppose it if a rich enough language existed. (It is, perhaps, part of the critic's role to enrich our linguistic resources in precisely this area.) The concept of specification resembles the move from 'coloured' to 'red'. If an object is reported to be coloured, a question is automatically raised as to what colour it is. This would not ordinarily be regarded as a request for the identification of an object, say redness, to which the coloured object stands in some relation.

The concept of expression in art can be clarified by applying to it the treatment Goodman applies to the representation of non-existent objects (Goodman (op. cit. pp. 21f.)). Representation, as it occurs in portraiture for example, is a two-place concept. A portrait is a representation of its model. But many representational paintings, such as pictures of unicorns, do not have a model. Goodman suggests that a complex single place predicate should be introduced to deal with these cases. A painting of a unicorn is a unicorn-representing picture. This specifies the kind of picture it is, not the thing it is a picture of. No reference is made to unicorns. Similarly, a painting may be a sadness-expressing picture and a smile may be a sympathy-expressing smile. These descriptions specify the kind of picture and the kind of smile without referring to emotions. The task for a theory of artistic expression, then, is very different from the task for a theory of behavioural expression as initially perceived. The aim of the latter was to describe the relation between behaviour and emotion when the first expresses the second. The aim of the former is to formulate the conditions under which it is proper to ascribe expressive properties to objects, to compare ascriptions of expressive properties to ascriptions of non-expressive properties and to provide an adequate account of how the expressive properties of an object are related to its non-expressive properties.

4.2 The unanswered question at the end of 2.2 was whether the relational theory of behavioural expression could usefully be supplemented by something like the property theory of expression applicable primarily, but not exclusively, to art. The examples given in Section 2 were such that (ii), William's expressing resentment by staying away for three days, was adequately accounted for by a purely relational concept of expression, if (ii) exemplifies any useful concept of expression at all. (iii), John's expressing sympathy by smiling, was left totally unexplained insofar as John may or may not have been feeling sympathetic at the time. (i) was partly dealt with by the relational concept, but appeared to require further explication because of the similarity intuitively discerned between (i) and (iii). The difficulty encountered by the relational theory led to pessimism about the prospects of an integrated theory of expression sufficiently powerful to deal with examples of all three kinds. But a straightforward combination of the relational and the property concepts of expression may provide a solution. The property theory will apply to cases of expression, artistic or otherwise, which are inaccessible to the relation theory. Thus (iii) will be interpreted as the explicit ascription of a physiognomic or expressive property to John's behaviour. (ii) is purely relational and (i) will be rendered as a conjunction, one of whose conjuncts will assert that John has a feeling of sympathy, while the other attributes the physiognomic property of sympathy to John's behaviour. 'John expressed his sympathy by smiling' becomes : 'John was sympathetic and John's smile was sympathy-expressing'. This would allow the relational concept to be dropped altogether, at least for the purposes of facial expression. It is applicable only to cases like (ii) which do not constitute a fundamental concern of this thesis. The property theory will account for (i) and (iii) in such a way that (i) is equivalent to (iii) and an assertion that John feels sympathetic. More formally, where 'x' is a variable ranging over persons, 'a' ranges over actions, 'G' is a predicate variable ranging over expressive properties and 'F' over emotional properties or states, the following structures correspond to (i) and (iii) :

(i) $\exists x \exists a : \text{Action}(x,a) \ \& \ Ga \ \& \ Fx$ (alternatively : Possesses (x, F-ness)).

(iii) $\exists x \exists a : \text{Action}(x,a) \ \& \ Ga$.

This, of course, leaves open the relation between predicates 'G' and 'F', which is a matter of semantics.

But this interpretation of (i) is plainly incomplete, since the original sentence certainly seemed to indicate a relation between John's smiling and his feeling of sympathy. Should we then simply tack on an additional clause : r(F-ness, a) and leave open the nature of r, which is to be determined by a general theory about the relation between mental events and actions? Or should the concept of expression as a relation be restored as a potential (to be filled out) definition of r? Neither proposal is very promising. The first fails to account for the fact that in (i), the term 'expresses' seems to relate John's smiling and John's sympathy. The second implausibly represents (i) as the combined operation of two otherwise autonomous concepts of expression. (iii) raises difficulties of a complementary nature. (iii') has the advantage of cancelling the apparent implications about the agent's emotions by rendering 'expresses F-ness' as a one-place predicate, 'G'. But the motive for cancelling that implication is that 'a' may denote an insincere expressive action. This is not strictly comparable to the motives which encourage descriptivism in art. In the former case, but not in the latter, there is something wrong or unusual in the agent's not feeling what he expresses. Such behaviour is insincere and ought to be understood as an instance of a kind of behaviour most of whose other instances are sincere. The theory of artistic expression extends straightforwardly to the fictitious expressive behaviour of the actor on stage. But it is as wrong to assimilate insincere to fictitious expression as it is to assimilate lies to novels. A unitary concept of expression must be found which is neither relational, in the existentially quantifiable token-token sense considered until now, nor predicative in an unstructured way. At the same time, the concept must do justice to the relational and predicative elements in expression.

5. Expressing that ...

5.1 At the beginning of his discussion of expression, Goodman introduces a distinction which superficially resembles and has the same motive as other distinctions between relational and non-relational expression.

That a person expresses sadness may mean that he expresses the feeling of sadness or that he expresses his having of that feeling ... a person may express sadness he neither has nor claims to have, or may have or claim to have a feeling he does not express. (Goodman (1976) p. 45)

This passage runs together two distinctions which, once separated and combined with earlier suggestions, will provide the material for a rudimentary three-way distinction. One of the concepts thus isolated is more promisingly applicable to facial expression than anything considered hitherto. In the phrase, 'his having of a feeling', 'his' can be understood in two ways. If it is treated as a literal possessive pronoun as in §2.2, this may yield 'the having of a feeling which he has' as an interpretation of Goodman's phrase, thus generating the same pattern of entailments as the earlier preoccupation with the expression by a person of his feelings. But Goodman's phrase can also be interpreted : 'the having by him of a feeling'. This is a new suggestion. A person may express the having by him of a feeling of sadness. There is no obvious implication that the person is sad.

In the gloss on his distinction Goodman again runs together two points. A person may express sadness he does not feel (have). That is one point. But a person may also express sadness he does not claim to feel. This is a separate point. The artist, when he expresses a feeling in his work, may not have and does not, solely in virtue of what his work expresses, claim to have that feeling. But there is a sense in which a person who expresses sadness in his behaviour does thereby claim to be sad - whether or not he genuinely is. Linguistically there are a variety of ways of expressing sadness, one of which is simply to say, 'I am sad'. An utterance of this sentence is usually an assertion by the utterer that the utterer is sad. However, it is not, on that account, never an expression of sadness (cf. Alston (1965) who criticises the expressive/assertive dichotomy as it was understood by the Emotivists). I submit that some behavioural expression contains a non-verbal but assertoric element of this kind and I suspect that this is partly what Goodman had in mind. A happy smile is a kind of non-verbal claim or assertion that the person smiling is happy. It may be true or false because smiles may be sincere or insincere.

It is clear that Goodman's distinction does not correspond to Tormey's distinction between an inferential and a non-inferential (relational and predicative) concept of expression. It is considerably more sophisticated. If there is a use for Tormey's inferential concept then we now have three kinds of expression, ie. three kinds of statement that can be made using the word 'expression'. First there are reports, like (ii) which entail that what is expressed is

genuinely felt. Secondly, there are reports like (iii) which do not carry this entailment but do involve the attribution to the agent of something like a claim to feel what is expressed. Finally there are descriptions of artistic and fictitious (histrionic) expression, which carry no implications either about what the artist or actor genuinely feels, or about how he claims to feel. It is the second kind of description that applies to facial expression.

The natural way to represent Goodman's notion that one can express one's having of a feeling in a familiar form would be to treat 'expresses' as a sentence-operator. This would give the following structure :

(iv) x 's ϕ -ing expresses that (Fx).

Again, it is immaterial whether the sentence operated on is represented as 'Fx' (eg. 'John is (feels) sympathetic') or as 'Has (x , F-ness)' (eg. 'John has a feeling of sympathy'). 'John's smiling expresses that John is sympathetic' will now translate the (iii) example. The (i) example will become : 'John is sympathetic and John's smiling expresses that John is sympathetic'. (iii') and (i') in Section 4.2 will now become :

(iii^{*}) $\exists x \exists a$: Action (x, a) & a expresses that (Fx).

(i^{*}) $\exists x \exists a$: Action (x, a), Fx & a expresses that (Fx).

The relation between the two conjuncts, 'Fx' and 'a expresses that Fx' is now far less obscure than it was. The conjunction is equivalent to the statement that the sentence operated on by 'expresses that' is true. In general sentences that assert the performance of sincere expressive behaviour will say of some action, a , that : a expresses that p and p is true - where p is some proposition about the agent's emotional or mental state.

The proposal may well fail to account for cases like (ii) and for artistic expression. If there is a concept of expression which is such that use of it entails the existence of what is expressed then it could still be represented as a sentence operator with the added stipulation that 'expresses' in this sense belongs to a class of verbs which includes 'reveals that', 'shows that', and so on, which is such that if ϕ is a member of it, ' $x \phi$ -s that p ' entails the truth of its embedded sentence. (cf. Unger (1972) p. 301, McDowell (1980) p. 126.) This would then just be a semantically different concept

of expression from that which applies to smiling and which, like 'means that', 'is prima-facie evidence that', does not entail the truth of the sentence it embeds. This conforms to our intuition that (ii) does indeed exemplify a different concept of expression from the one which is exemplified in both (i) and (iii). Similarly, artistic expression just is different from facial expression, so the failure of my proposal to account for the former is not an objection to it. All that is required is an integrated theory that works for cases like (i) and (iii). If, furthermore, it distinguishes these from (ii) and artistic expression, then this is an advantage, not a disadvantage.

5.2 However, there remains a difficulty with this proposal. Construed as a sentence operator, 'expresses that' closely resembles 'says that'. But many would find the following form of 'says that' statement unacceptable :

$\exists x : x \text{ says that } Fx.$

The objection to the proposal in §5.1 is that, like sentences in the above form, it involves quantifying into a referentially opaque context from outside. There are three possible responses to this objection. One might first attempt to answer it by pointing out that the objection presupposes that the presence of 'that' invariably introduces opacity in the subject position of the embedded sentence. Since the pre-supposition is false, the objection could be met by treating 'expresses that' as an addition to a class of operators, which includes 'it is true that ... ', whose members do not interfere with reference. Alternatively, we could concede that there is opacity and proceed to deal with it by applying to 'expresses that' statements whatever method is preferred for the analysis of opaque sentences in general, eg. Davidson's 'paratactic' analysis or Quine's (1966) proposals.

I do not propose to choose between these strategies or to describe either in detail because the issue of opacity is, in the present context, of negligible importance. Even if it is conceded that 'expresses that' is opaque, the main concern of this thesis is with comparatively primitive kinds of expressive behaviour, kinds of activity in which not only humans but also unsophisticated, languageless creatures can participate. This is relevant because if replacements

for x in 'x expressed that (Fx)' are to include primitive creatures then ordinary inferences which assume the applicability of Leibniz' Law to the second occurrence of x will, in any conceivable case, derive true conclusions from true premisses even if they are not formally valid. This is true not just for primitive creatures, but for sophisticated ones using primitive means. Thus if it is true that 'John expressed that John was sympathetic', and John turns out to be Mr Smith, then it is difficult to imagine circumstances under which the results of substituting 'Mr Smith' for either or both occurrences of 'John' would be false.

5.3 Finally, something should be said about the way in which applying the 'expresses that' concept to physiognomic behaviour allows the element of truth in the relational theory to be preserved. For, while persisting in the denial that the relevant concept is to be defined as a relation between an emotion and a piece of behaviour, it now seems that what distinguishes behavioural from artistic expression is that the former involves a relation between an expressive agent and his behaviour which is not duplicated by the latter. Behavioural expression *attributes* what is expressed to the agent. No such attribution characterizes artistic expression. The point can be put in quasi-semantic terms : expressive behaviour "refers" to its agent. Artistic expression, which might be supposed to carry a certain aspiration towards universality, does not "refer" to the artist. True, this insistence on self-reference, while it allows for the coherence of insincere expression, imposes a certain complexity on the analysis of its histrionic counterpart. The complexity will match that of analyses put forward by Tormey and others (Tormey (1971) p. 52). It is misleading to say that Olivier expressed Othello's jealousy. Rather, we should say that Olivier portrayed Othello expressing jealousy.¹⁰ This is not, as Tormey thought, because the relevant concept of expression asserts a relation between an emotion and the expressive product, but because it asserts a relation between the product and its author which significantly transcends the given relation of authorship or production.

10. Histrionic expression, therefore, like fictitious as opposed to poetic expression, is not strictly comparable to aesthetic expression in general, where that is understood to apply only where we say that the work of art expresses F-ness. A better comparison is between histrionic/fictitious expression and cases where a painting is said to represent a person expressing F-ness.

This concludes my survey of the grammar of the concept of expression in which I am interested. For the sake of verbal facility, however, I will continue to use idiomatic constructions in the rest of this essay. The assumption is that the sense of my usage will conform to the requirements of this chapter despite its idiomatic surface grammar.

II. INTERPRETATION AND EXPLANATION

1. Utterances

1.1 To interpret an utterance is to assign a meaning to it. The same utterance, given the truth of assumptions about the speaker's truthfulness and literalness, is partially explained by the attribution of a relevant belief to the speaker. There is an obvious parallel between this distinction and the contrasting conceptions of expression discussed in Chapter 1. To explain an utterance, or any other kind of action, is to indicate how it is (causally) related to the beliefs and desires of the speaker or agent. To interpret an utterance is to say in some detail what kind of action it is without saying anything about its causation. The effect of the previous chapter was to supply a purely formal adumbration of a hermeneutic rather than an explanatory construal of descriptions of behaviour as expressive. But this ostensibly perspicuous distinction is repeatedly threatened. The source of the menace lies not only in what appear to be explicitly reductive programmes in the philosophy of language - the traditional target of this kind of criticism being, perhaps unfairly, the Gricean programme - but also unexamined confusions pertaining to non-linguistic theories of interpretation. Confusions of this kind are prevalent in theories about action, especially in the philosophy of the social sciences, where two paths can be discerned which lead to the same predicament. First there is the hermeneutic tradition, traceable to Dilthey, which is based on the model of textual exegesis.¹ In its cruder manifestations, this school of thought seeks to replace psychological explanation altogether with a species of interpretation. Less crudely, its exponents still purport to have isolated an interpretative element, or presupposition in all psychological explanation. More recently, some philosophers, fascinated by the work of Davidson and Quine, have focused exclusively on the interpretation of behaviour.² A failure to distinguish interpretation from explanation also obscures certain questions about perception, where a clarification of the distinction will yield results interestingly germane to the theme of

1. See, for example, Ricoeur (1971).

2. eg. Hookway (1978).

behavioural expression. The present chapter will proceed via some initial remarks about radical interpretation to a consideration of the extent to which similar techniques are applicable to non-verbal behaviour and an attempt to disambiguate a class of action-descriptions by reference to the interpretation/explanation dichotomy. Finally, a parallel dichotomy is exposed which applies to perception.

1.2 There are features of Davidson's (1973, 1974) writings on radical interpretation which may partly explain the conflation of interpretation and explanation by subsequent writers. It must be stressed, however, that no confusion on this score is imputable to Davidson himself. The latter has discerned two analogies and two connections between explanation and interpretation; both the analogies and the connections fall short of equivalence. First, the former. To interpret a German speaker's utterance of 'Es regnet' is to re-describe it as an action of saying that it is raining (Davidson (1973) p. 314 & (1974) p. 309). Decision theory, the function of which is to explain people's actions as the product of their desires and beliefs, also 'allows us to redescribe certain events in a revealing way'. But Davidson explicitly warns against taking the analogy any further, at least in one inviting direction.

For decision theory can also explain actions, while it is not at all clear how a theory of interpretation can explain a speaker's uttering the words 'Snow is white'. ((1975) p. 13)

A theory of interpretation will yield a different kind of information from that yielded by an explanatory theory such as decision-theory, even where both types of theory are brought to bear on the same material. The analogy, therefore, is strictly incomplete.

Nevertheless, there are finer structural features which the two kinds of theory share. The organisation of the conceptual material required for radical interpretation maps onto that of the material required in the explanation of choices. In the latter, the agent's subjective probability and the relative strength of his desire conspire to determine his preferences vis-a-vis a specific context of choice. His preference then unilaterally determines his choice. In the explanation of speech, the speaker's belief and the meanings he attaches to the sentences of his language conspire to determine which sentences he holds true at any particular time. It is to be assumed that evidence that a speaker holds a certain sentence to be true is unproblematically accessible in advance of knowledge both of the speaker's

beliefs and of the meaning of his utterances (Davidson (1974) p. 312-313). Holding-true unilaterally explains the speaker's linguistic behaviour, at least within an extensive, assertoric fragment of his language; it determines which sentences he will produce when interrogated and which sentences uttered by others he will assent to. Utterance, holding-true, belief and meaning in radical interpretation map onto choice, preference, subjective probability and strength of desire, respectively, in decision theory (ibid. p. 314).

The two connections between interpretation and explanation are partially implicit in the structural analogy. The first concerns the need, in experimental choice situations, to establish a common language uniting the subject of the experiment and the investigator (ibid. p. 315). This need not preoccupy us here and I mention it only in order to distinguish it from a second connection which is more germane to the present theme. As Davidson understands it, the radical interpreter's problem consists in the impossibility of determining the beliefs of members of a community under investigation before he possesses a theory to interpret their utterances, but such a theory is not available in the absence of prior information about their beliefs. Consequently, 'we must somehow deliver simultaneously a theory of belief and a theory of meaning' (ibid. p. 312). The function of the holding-true attitude in the theory is to show how this is possible. It is assumed that the conditions under which native speakers hold sentences of their language to be true are accessible to the interpreter. Instances of holding-true, therefore, can be substituted for utterances as the evidential base, the explananda of the theory. The fact that speakers hold a sentence to be true under specifiable conditions will provide evidence simultaneously for the ascription of a belief to speakers under those conditions and for the assignment of meaning (truth-condition) to the sentence.

What does this imply about the connection between meaning and belief, between interpretation and the doxastic component in the teleological explanation of speech-acts? There is a puzzle here. The procedure which allows us to interpret radically alien linguistic performances is the same as the procedure which allows us to ascribe beliefs to the speakers of radically alien languages. Both are a matter of correlating linguistic behaviour with 'stimulus conditions', where linguistic behaviour includes utterances as well as assent to, or dissent from utterances by others. This seems to suggest that the

connection between interpretation and belief ascription is not mere simultaneity, but equivalence. How could there be a difference of substance between two theories with a shared evidential base? The question is, of course, manifestly verificationist in tone, and one could, in the spirit of Quine, answer that there is no difference. Quine, indeed, does not always distinguish between questions about the reality of the mental and questions about the determination of semantic facts (eg. Quine (1969) p. 27f). But this would be alien to the spirit of Davidson's philosophy, which includes a commitment to an ontology of mental events (Davidson (1970)). There seem to be two options here. Either one can retain a commitment to a form of behaviourism, but hold that there can be behavioural evidence for belief which is independent of language. This is the strategy adopted by Jonathan Bennett in (1976). Or it can be denied that the dependence of two theories on the same evidence entails the equivalence of the theories with respect to their substance.

Davidson's position conforms to the second alternative. It is not meant to justify a reduction of thought to speech. The fact that interpretation and belief-ascription share the same evidential base is not explained by their theoretical equivalence but by the reciprocal nature of the evidential links between them. '... is evidence that ...' is a transitive relation between sentences. If p is evidence that q, then whatever evidence there is for p is also evidence in support of q. If, in addition, q is evidence that p, then whatever evidence there is for q is also evidence in support of p. This describes the situation in radical interpretation, when p might be a sentence assigning a meaning or truth-condition to a sentence of the object language and q a sentence ascribing a belief to the utterer of the sentence. There is no temptation to suppose that p and q are in any sense equivalent.

Davidson's own conception of his theory is definitely non-reductive ((1975) p. 9). The connection between interpretation and first-order belief-ascription is displayed in a description of the role of interpretation in the teleological explanation specifically of utterances. The passage quoted on p. 29 continues as follows :

'... uttering words is an action, and so must draw for its teleological explanation on beliefs and desires. Interpretation is not irrelevant to the teleological explanation of speech, since to explain why someone said something we need to know, among other things, his own interpretation of what

he said, that is, what he believes his words mean in the circumstances under which he speaks. Naturally this will involve some of his beliefs about how others will interpret his words. ((1975) p. 13)

In general, the explanation of action requires information about what kind of action the agent takes his behaviour to be. In the case of speech this primary information is provided by the theory of interpretation. The product of the theory of belief is not the same as the product of the theory of interpretation. The first picks out the speaker's first-order beliefs about the world. The second picks out his second-order beliefs about how others will interpret his behaviour.

1.3 It will be pertinent at this point to mention some of the difficulties which afflict the Davidsonian conception of radical interpretation. The nature of the criticisms to follow is not such as to entail rejection of the theory on the ground that it cannot avoid reduction. The point is rather that certain conspicuous silences in the Davidsonian theory seem to encourage the search for a supplementary or even an alternative theory elsewhere. The obvious direction in which this search will lead is towards the theory of communication-intentions and this, notoriously, introduces a new risk of reduction. For the Gricean programme appears to construe interpretation as the explanation of an utterance as the outcome of a complex set of intentions vis-a-vis an audience. This kind of programme is not, in fact, as pernicious as the last sentence makes it appear. This is seen once it is supplemented by a distinction between the interpretation of token-utterances and the interpretation of utterance-types (cf. Schiffer (1972) p. 3f).

A theory of interpretation is bound to assume the existence of a high correlation between the beliefs actually held by speakers of a language and the beliefs they express in their linguistic behaviour. It is not, therefore, an objection to Davidson that his account is firmly based on this assumption. But mental events in general cannot, given 'the holism of the mental realm' (Davidson (1970) p. 217 in (1980)), be attributed singly.

Beliefs and desires issue in behaviour only as modified and mediated by further beliefs and desires, attitudes and attendings, without limit. (ibid.)

The ascription of a belief to an agent will explain his behaviour, verbal or otherwise, only in conjunction with a specific desire. If there is reason to believe that the relevant desire is absent and that

some other desire is operative, then the same action will require for the completion of its explanation the ascription of an altogether different belief to the agent. It is a corollary of this that if instances of a certain kind of behaviour are regularly associated with the same belief, then they must also be associated with a whole set of beliefs and desires. The objection to Davidson's account of radical interpretation is now seen to be not that it imports too much explanatory material into the theory of interpretation, but that it contains too little. If Davidson's description of holism is literally true, then it seems that the amount of explanation required before interpretation can begin is limitless. But the limitlessness of the antecedents of action is, I think, an exaggeration. In many cases little more is needed than pairs of desires and beliefs (cf. Bennett (1976) p. 80), and in all, explanation 'must come to an end somewhere'. The relevance of many of an agent's attitudes to particular actions will be highly attenuated. What, then, is required for the explanation of speech beyond the attribution of the belief expressed? At the very least the speaker must want his interlocutor to be correctly informed. The radical interpreter must assume a certain complicity on the part of his informer. And Davidson himself has conceded (in the passage quoted on p. 31) that the speaker must have certain beliefs about how his words will be interpreted by others. It is difficult to see, in the light of these remarks, what objections remain to the incorporation of the notion of communication-intentions into the interpreter's theoretical apparatus.

A similar conclusion can be reached by a slightly different route. The argument of the previous paragraph is compatible with the view that no evidence for a person's beliefs is obtainable in advance of an adequate interpretation of his utterances. The argument to follow puts that claim in doubt. It could be argued that all that Davidson has shown is that, in Peacocke's (1979) phrase, no belief can be attributed to a person in advance of the application of other concepts belonging to some scheme of holistic explanation. Radical interpretation is not the only such scheme in which the concept of belief has an explanatory function, the other being the scheme of action-explanation. A defender of Davidson could offer two answers to this challenge. The first answer would depend on the first-mentioned connection perceived by Davidson as holding between interpretation and explanation. In experimental situations devised to test a subject's preferences between options, the values of which are not straight-

forwardly discriminable, the experimenter must be able to check that the subject correctly understands his instructions (Davidson (1974) p. 315). This raises anew the problems first raised in connection with interpretation alone. Ideally, therefore, one must aim for an integrated theory of decision and interpretation (*ibid.* p. 316), and the interpretation and action schemes should be seen to constitute a single scheme of holistic explanation. However, the suggestion that all preferences are inaccessible in the way in which they are in the experiment discussed by Davidson is implausible. In general, where there is a choice between two options which are manifestly different, preference is straightforwardly exhibited in the choices made. Of course, this does not settle which of several possible explanations of the preference exhibited is true. This is still a matter of attributing sets of beliefs and desires to the subject, but it was not at this point that Davidson saw a need for interpretation, and it is highly unlikely that there is a specific need here (cf. Bennett (1976) Chapters 3 and 4). Precisely which set of attitudes will truly explain an agent's activity on a particular occasion is a question the answer to which is determined by the agent's behaviour in different but comparable contexts. This may include his linguistic behaviour but need not do so.

It will not help Davidson's case to point out that these objections entirely ignore the role of the holding-true relation in the theory. For an additional difficulty in the theory concerns the explanatory weight placed on that concept. The burden is more than it can bear for the following reasons. First, the concept of holding-true is itself an explanatory concept which is impotent outside an appropriate scheme of holistic explanation.³ It is true that we can sometimes tell whether a man holds a sentence to be true without knowing either what belief of his the sentence expresses or what the sentence means. But Peacocke ((1979(p. 199) has argued that this ability presupposes that applications of the entire scheme have had successful results with the same person in the past. If, for example, we can tell that a speaker is being honest from his tone of voice,

3. In developing his theory of holistic explanation Peacocke argues that Davidson's characterisation of the holism of the mental realm fails to capture the distinctive nature of holistic explanation in general (Peacocke (1979) p. 20). But the arguments of this section require only that Davidson's characterisation should be true of the mental realm, which Peacocke allows. The other details of Peacocke's highly intricate theory need not detain us here.

this can only be because we have been able in the past correctly to attribute beliefs to him on the basis of his utterances in a similar tone, of sentences we do understand. This is plainly unlike the situation in radical interpretation. The impasse which faced interpretation at the outset has not been removed. This particular objection could be met if there were evidence of a different kind to support the claim that there are natural signs of honesty, the recognition of which is not impeded by cultural barriers. The nature of the evidence required is discussed in the final chapters of this thesis.

However, if such evidence were available, it would not just be evidence in favour of the holding-true attitude, but would have to support the simultaneous ascription of a much more complex set of attitudes to the speaker. For holding-true simply does not determine linguistic behaviour unilaterally in the way in which preference determines choice. This is not just because of the obvious point that at any one time there are several sentences a person holds-true, most, if not all, of which he does not voice. The relevant linguistic behaviour does not only include a speaker's spontaneous utterances but covers the sentences produced by the speaker under interrogation and the sentences to which he assents. Such behaviour is still not exhaustively explained by the speaker's holding the sentence in question to be true. For he will not produce or assent to an appropriately true sentence of his language unless he is predisposed to supply his interlocutor with correct information about his present beliefs and, in the context of radical interpretation, about the meanings of sentences in his language. It is here that a proponent of the Gricean theory might be expected to discern a suitable aperture through which to infiltrate his theory. A crucial component of that theory has already been admitted, albeit in an imprecise form. The notion in question is that of the speaker's attitude towards the interpreter, mention of which is required to supplement the attitudes towards sentences already mentioned in the explanation of linguistic behaviour. It is deeply plausible to argue that the nature of the native speaker's complicity is best captured by a description of the speaker's intentions vis-a-vis the interpreter. Remaining within the context of indicative utterances, the speaker's attitude will be characterised as an intention to produce or 'activate' a true belief in the interpreter about what he, the speaker, believes at the time of utterance; and this intention will, itself, be intended by the speaker to be 'overt',

ie. recognised by the interpreter.⁴ The aim of the theory will then be to specify the extent of the overtness required, that is, the extent of the knowledge the interpreter must possess if he is to interpret the speaker's utterances.

1.4 This development introduces a deeper threat to the interpretation/explanation dichotomy which it is the aim of this discussion to preserve. This new threat is, in a sense, the opposite of that which was discerned in a behaviourist concept of radical translation at p. 30 above. There it began to look as if the concept of belief could be made redundant in the explanation of speech. It seemed to add nothing to what has already been achieved when meanings have been assigned to utterances. For completion of the latter task is reached when speakers' dispositions to assent to sentences of their language have been correlated with observable conditions. To maintain that speakers believe what they say is to contribute nothing further to the explanation of their linguistic behaviour since belief, too, is exhibited only in observable disposition to assent to sentences. The communication-intention theory reverses this tendency. Now it seems that the attitudes of the speaker to the world (beliefs) and to an audience are sufficient to explain his communicative behaviour. Nothing further is contributed to the explanation by an independent notion of sentence-meaning. The explanation of an utterance on the basis of the utterer's intentions constitutes its interpretation.

As thus summarised, however, this style of theory has little chance of providing a complete account of linguistic meaning. The theory is construed as a straightforward reductive proposal involving a single step from the concept of what a speaker means by what he utters on a particular occasion, which is defined in terms of his intentions vis-a-vis an audience, to the meaning of his utterance. In this form the theory is vulnerable to the objection, raised by Mark Platts ((1979) pp. 88-90), that it makes no mention of the belief held by the utterer of a sentence, σ , which fulfils the essential role of justifying or rationalising the utterance of σ by a speaker who intends

4. This formulation is intended to be approximate since the text would be uselessly cluttered by a detailed exposition of the analysis of the speaker's intentions. It follows Strawson ((1970) at p. 172 in (1971)), with the modified notion of the 'activation' of belief imported from Schiffer ((1972) p. 51) and the stipulation that the belief intended by the speaker to be activated in the interpreter be true, imported from McDowell ((1980) p. 125f).

to communicate that p . This belief, so the objection continues, can only be specified as the belief that σ means that p . But this objection is effective, if at all, only as directed against a theory produced by an extraordinarily uncharitable interpretation of Grice's earliest (1957) article (p. 46 in Strawson ed. (1967)), and also against Bennett's conception of 'one-off' cases of S-meaning, insofar as the latter, as initially introduced, is explicitly meant to be independent of any behavioural regularity and no explicit reference is made to an alternative semantically explanatory concept, such as iconicity (Bennett (1967) p. 13).

The programme of the communication-intention theorists, however, constitutes a reduction which proceeds via at least two stages. It cannot proceed without the introduction of the concept of an utterance-type as distinct from that of an utterance-token. The order of logical priorities, for the communication-intention theorist, runs from the concept of what a speaker means by an utterance on a particular occasion through the concept of the meaning of a whole-utterance type to the concept of the meaning of a whole-utterance token (Schiffer (1972) p. 7). A whole-utterance type, σ , means that p only if that is what people generally mean when they utter tokens of σ . The concept of speaker-meaning, then, enters only indirectly into the definition of the concept of the meaning of a whole-utterance token. Mark Platts' objection is met, therefore, by agreeing that what justifies a speaker's belief that his intentions will be realised is the meaning of the utterance type, a token of which he utters. This will then be analysed on the basis of the complex intentions with which tokens of that type are normally uttered. The account still depends, of course, on there being cases of S-meaning which are successful but do not depend on language. That cases of this kind are possible ought not to be seriously disputed, but no doubt a reductive programme of the Gricean variety needs to say something about how they are effected. Such an account will need to appeal to some broad concept of 'iconicity' (cf. Bennett (*ibid.*) p. 139f) or 'salience' (cf. Lewis (1979) p. 35f). At this point the details need not concern us further. The relevant consequence of the preceding precis of the Gricean programme is that even if that programme is reductively construed, it still admits a distinction between interpretation and explanation in relation to utterance tokens. To interpret an utterance is to specify the type of which it is a token. To explain an utterance is to attribute a complex set of beliefs and intentions to the speaker. However, there remain

grounds for dissatisfaction with this proposal.

2. Facial expression and the linguistic analogy

2.1 In insisting on the difference between interpreting and explaining utterances it has not been necessary to introduce the possibility of a systematic discrepancy between the beliefs held and the beliefs expressed by the speakers of a language. To countenance such a hypothesis is to sabotage radical interpretation which can be initiated only on the basis of an assumption that there is a reasonable correlation between belief and speech. This assumption is not purely methodological. The predicament of the radical interpreter is comparable to that of the child acquiring what is to become his native language. Unless considerable success in attributing beliefs on the basis of the linguistic behaviour of others is achieved it is difficult to see how a language could be learnt, how, that is, language could exist. But it seems that the initial assumption must include far more psychological material than Davidson was apparently prepared to allow. The right kind of supplementary material is supplied by the concept of S-meaning. Even if this is correct however, a distinction can still be preserved between explanation and interpretation, interpretation being defined as the standard explanation of utterances of a certain type.

Nevertheless, the correlation between belief and assertion, or between what a speaker means by uttering a sentence and what that sentence means is, though indispensable, far from perfect. A theory of interpretation cannot have as one of its consequences that people never assert what they believe to be false and never deny what they believe to be true. The assumption required by the interpreter corresponds rather to the *apriori* notion that dishonesty is possible only against the background of widespread honesty. This is recognisably a version of the principle of charity (cf. for example, Davidson ((1974) pp. 320-321); one which provides not only for a general reasonableness, but also for a general propensity to honesty. The imperfection of the correlation involved is a *prima-facie* reason for rejecting a reductive construal of radical interpretation, should one ever literally be suggested. The reduction of sentence-meaning to speaker-meaning might seem better off in this respect. But the prevalence of non-literal uses of language is an obstacle to a total reduction of this variety also. Metaphors are not always original. The sage cliché

which discourages putting all one's eggs into a single basket is rarely offered as advice about how to gather eggs. What people mean by their words may, for a minority of sentences, fail to tally with what their words mean. The justification for sustaining the idea that there is a contrast in such cases between literal and metaphorical meaning will appeal to the notion of semantic structure, the idea that there are antecedently established principles which determine the meanings of sentences on the basis of the meanings of their components.

Semantic structure, and the meanings it determines, must be allowed at least this degree of autonomy even from standard linguistic practice. I do not mention this point in order to revive the debate between theorists of formal semantics and theorists of communication-intentions. I doubt whether there is a substantial issue here.⁵ The role of truth-conditions in the statement of a theory of meaning for a language is not put in question by any of the earlier discussion of interpretation. Their admission here does not imply doubts about the relevance of speakers' attitudes. The point here is entirely free from polemical motivation. The more thoroughly the elementary principles of a natural language are entrenched (by which I mean whatever is described by the axioms of a semantic theory), the more its semantic potential outstrips its actual use and the less is it subject to modification under the influence of alterations in the linguistic practice of its speakers. Compatibly with this it is quite reasonable to maintain that what determines the correctness of the basic axioms is the way most speakers use the sub-sentential components they define in most of the sentences in which they use them. What is not acceptable is the assertion of a straightforward reductive link between whole-utterance types and linguistic practice for highly developed natural languages.

2.2 A potentially fruitful analogy with non-linguistic expression suggests itself at this point. The counterpart in the theory of non-linguistic expression to the Gricean account of linguistic meaning, is the idea that the meaning of (or what is expressed by), for example, laughter is determined by what people standardly feel when they laugh.

5. Cf. Strawson (1970), Peacocke (1976) and McDowell (1980). There is still room for debate, of course, about the exact nature of the contribution of structure and of speakers' attitudes, about how they collaborate and about how they diverge.

However, Wollheim ((1964) p. 284) suggests that it is by no means absurd to suppose that there are people who always laugh when they are sad without thereby expressing sadness. There could exist a form of 'pathognomic disturbance' such that, for a certain range of emotions, people invariably expressed the contrary of the emotion they felt. This raises the following question : what is it that contributes to the determination of the meaning of facial expressions what semantic structure contributes to the determination of the meaning of utterances? If there is no answer, then Wollheim's hypothesis must be dismissed as absurd. If there is an answer, it must specify an attribute of facial expression which is autonomous in relation to psychological determination to the degree to which semantic structure is so autonomous. Two proposals have been put forward which I discuss in Chapters 6 and 7 respectively. The first, and by far the most plausible, is Wollheim's own view that facial expression is characterised by a kind of appropriateness or matching of the expressans to the expressandum; an appropriateness, however, which obviously falls short of literal iconicity or resemblance. The second idea, which prevails in the anthropological literature, is that facial expression and gesture are determined by their own brand of semantic structure which is closely akin to the structural features of language. I do not pursue these questions here, leaving open for the time being the question of whether the interpretation of behaviour is wholly, partially or not at all determined by its standard explanation.

3. Describing Behaviour

3.1 How vague or trivial a doctrine is Behaviourism? I do not propose to offer a detailed evaluation here, but merely to draw attention to a certain confusion exhibited in some attempts to settle this question. It will emerge that clarification can best be achieved by applying the distinction between interpretation and explanation to the description of action. The challenge is to find some general statement of the relation between behaviour and its psychological antecedents which would appeal to behaviouristically-minded philosophers and which is not either uninterestingly true or plainly false. I do not believe that this challenge can be met, and this is not a new thought. The contrast between what follows and what has been written in the past concerns descriptions of behaviour which are such that the claim that they entail psychological descriptions has previously been dismissed

as true but uninteresting. I will argue that the claim could equally be regarded as interesting but on the whole false, depending on how the descriptions of behaviour are interpreted. There are two legitimate options.

Paul Ziff ((1958) p. 157 in (1966)) has shown that it is impossible to refute Behaviourism on the ground that there is a principled objection to the statement that one can always find out whether someone is angry but that the same objection does not bear on the claim that one can always find out whether someone is behaving in a certain way, eg. gnashing his teeth. Ziff's point is simply that one cannot in principle always find out whether someone is behaving in a certain way. True, one can establish by means of observation alone that someone's teeth are gnashing, but one cannot establish that someone is gnashing his teeth with the same facility. The difference is between an instance of a bodily movement in a transitive sense and an instance of a bodily movement in an intransitive sense. 'To gnash' belongs to a class of verbs, lucidly identified by Jennifer Hornsby ((1980) p. 2) which function transitively and intransitively and which are such that transitive uses of them entail intransitive uses but not vice-versa. 'John gnashed_T his teeth' entails 'John's teeth gnashed_I'. But the gnashing_I of John's teeth does not entail John's gnashing_T of them. Ziff's point, with which I agree, is that if behaviour consists in bodily movements_T then it cannot be known to occur by means of observation alone. For Behaviourism to be a substantial doctrine, it would have to identify behaviour with bodily movements_I, thus degenerating into blatant falsehood.

3.2 However, too much is apparently conceded to Behaviourism in the following remark :

I deny that there is a difference between finding out whether or not I am behaving in certain ways and finding out whether or not I am angry. (Ziff (ibid.) p. 160)

This seems to suggest that there is no difference between finding out whether or not John is gnashing_T his teeth and finding out whether or not John is angry. This is wrong because to find out whether John is angry is not merely to find out whether he is gnashing his teeth, but to find out why he is doing so. He may be doing so for some other reason - in order to look as if he is angry, or merely in order to exercise his jaw. But gnashing_T of teeth is not, perhaps, among the ways of behaving, the kinds of behaviour Ziff had in mind at this point. There is a class of descriptions of actions, members of which are

formed by prefixing the noun 'behaviour' with the name of a mental state or event : 'anger-behaviour', 'pain-behaviour', etc.. The first example occurs several times in Ziff's article, whose point, therefore, is that there is no difference between finding out whether I am anger-behaving and finding out whether I am angry. Is anger-behaviour the same as angry behaviour and are gnashings_T of teeth always, nearly always or only sometimes instances of it?

These questions signal an ambiguity in descriptions of the general form : 'F-ness-behaviour'. They could be used either to interpret or to explain the behaviour they denote. On an explanatory construal the passage quoted above will come out true. 'F-ness-behaviour' is equivalent to 'behaviour exhibited because the agent is F'. Finding out whether John is angry is almost the same as finding out whether he is acting in a certain way because he is angry - not quite, however, because one could conceivably establish that John is angry without establishing that this is why he behaves as he does. But this does not block the relevant entailment which runs from 'x is F-ness-behaving' to 'x is F'. But an interpretative construal will make the quoted passage come out false. As an interpreting description, 'anger-behaviour' is equivalent to 'anger-expressive behaviour', or 'angry behaviour'. Gnashings_T of teeth are sometimes, perhaps even nearly always, instances of anger-behaviour explanatorily construed. They are always instances of anger-behaviour interpretatively construed. More precisely, teeth-gnashing, scowling, fist-clenching, etc. are timeless kinds of anger-behaviour.

Behaviourism is not obviously false if it is the claim that there is no difference between finding out whether someone is in a certain psychological state and establishing that a certain interpretation of his behaviour is correct. An observer could stand in a relation to an anger-behaving subject which is akin to the relation in which the radical interpreter stands to the speaker of an uninterpreted language. This is the predicament of the anthropologist investigating radically alien communities and refusing to suppose that there are any cross-culturally constant patterns of behaviour (cf. Chapter 7). The problems of the anthropologist are comparable in certain ways to those of the field lexicographer. However, they are in other respects both less complex and more complex. The psychological explanation of much expressive behaviour may be less complex than that of rational intentional action including speech (see Appendix) being less than fully intentional. But ensuring a correct

description of eliciting stimuli may be more difficult (see Chapter 7). Given the first point, the obstacles to a reduction of comparatively simple psychological facts to facts about behaviour may seem to be diminished. But given the second point, the epistemological value of such a reduction will be negligible.

However, the familiar objection remains that people are not always honest or sincere. They disguise emotions they feel and simulate emotions they do not feel. A distinction, therefore, must be sustained between the interpretation and the explanation of instances of expressive behaviour. However strong the pragmatic case for a purely methodological behaviourism as a basis of the anthropological attitude - an attitude which can be adopted at home just as radical interpretation can be practised on our native tongue - a conceptual behaviourism is not justified. Equally cogent practical reasons tell against the latter, a strict adherent to which would leave himself open to perpetual deception. There is, I submit, some difference between finding out whether our neighbours are anger-behaving and finding out whether they are angry. The difference is partially exhibited in the common-place fact that we do not ordinarily register a need to investigate at length what people's behaviour means, but we are frequently conscious of such an exigency with respect to what people mean by their behaviour.

3.3 I indicated (on p. 28) that one consequence of the failure to discriminate between interpretation and explanation is that the explanation of action is mistakenly taken for a species of interpretation. The less extreme version of the hermeneutic proposal might still appear to be compatible with what has been said so far. The weaker hermeneutic claim may be put thus : just as speech-acts cannot be explained without antecedently or simultaneously being interpreted, so all forms of human behaviour require interpretation as a stage or component of their explanation. I believe that we are now equipped to show why even the weaker hermeneutic position is wrong, or at least highly implausible. The concepts of interpretation and expression are complementary with respect to non-linguistic behaviour as are the concepts of interpretation and meaning in connection with language. It is the contention here that (pace Charles Taylor) not all actions are expressions, natural or otherwise, of 'tryings to get' (Taylor (1971) p. 81), and the idea that only expressive behaviour is a proper subject for interpretation is sufficient to settle this in advance of

any decision about the precise nature of interpretation. If interpretation depends on some notion of appropriateness or iconicity then an immense range of actions are excluded. True, an action will be appropriate as a means to a desired end : but it will not match the desire in the relevant sense. Smiling, however, is felt to match happiness in some way (Wollheim (1966) p. 99 : (1968) p. 48 in 1970 edition). If interpretation is characterised in terms of standard explanation, then the range of behaviour admitted as expressive will be somewhat wider. It would have to include, for example, eating as an expression of hunger since eating is standardly explained by hunger. But much is still excluded in virtue of the fact that most actions, qua bodily movements_T, do not have standard explanations. (cf. Chapter 6, §3.4)

4. Describing perceptual experiences

4.1 There is an analogue to the interpretation/explanation dichotomy in the theory of perception, and there is a striking parallel between the tendencies introduced by its neglect here and those introduced by its neglect in the theories of meaning and action. In the latter areas failure to observe the distinction gives rise to a form of behaviourism. In the context of perception, the species of reductionism which conflation encourages is phenomenalism, initially, but in speaking of the perception of human behaviour, the risk of behaviourism is repeated. The source of the problem is an ambiguity in standard formulations of a fundamental question concerning the role of inference in perception. The question is about the nature of the cognitive act or acts which constitute perception of objects - is it purely a matter of the operation of the senses, or does perception require the intervention of a further, more intellectual cognitive faculty? In Chapter 5 I pursue a comparable question raised about the recognition or perception of expressive behaviour. At this point, the question is left unanswered while an attempt is made to formulate it more perspicuously. The statement of the issue is silent as to the nature of the cognitive product of the act in question, and it is at this point that the distinction between interpretation and explanation reappears.

Two questions must be distinguished accordingly. First, for any cognitive act we must ask : how does it provide its subject with an interpretation of its raw input? Applied to the issue of perception the questions are : how much intellectual work is demanded from

the perceiving subject if he is to have an experience *as of* there being some physical object, eg. a chair, before him? and : how much intellectual work is demanded from the perceiving subject if he is to have an experience *of* a physical object, eg. a chair, before him? I may as well declare my prejudice in favour of one answer rather than another. I believe that no intelligence is required for the perception of physical objects or for the having of experiences as of physical objects of fairly elementary kinds. First, any visual impression that is appropriately (non-deviantly) caused by a physical object, whether or not the description of the former fits the description of the latter, is a perception. Furthermore, the account of its causation will not mention any propositional attitudes. Secondly, any experience, hallucinatory or veridical, as of a physical object will be interpreted to a degree independantly of the operation of some other cognitive faculty. All seeing is seeing-as (as the following, among others, have said : Howell (1972) : Ishiguro (1967) : Vesey (1956)) and every illusion is an illusion that an object of a certain kind is present. In each case the actual or putative object is at least seen as having a certain colour, shape and size. (To the extent that they are relevant, these themes are developed in Chapter 5.)

4.2 Radical interpretation has its analogue in perception. The situation can arise where we neither know how to interpret nor how to explain a certain experience. It is clear how this could be the case for someone else's experience. I may want to know what someone else can see but be ignorant of what lies within his field of vision. The situation is such that I neither know which object or objects are being perceived nor what they are being perceived as. Both questions can be answered simultaneously by means of an elementary description of an object in terms of its basic perceptual properties. The evidence sought might consist in the perceiving subject's response to an appropriate question. In normal conditions, evidence which supports the statement : 'x sees an F' will also support the statement : 'There is something which x sees as an F' or 'x is having an experience as of an F'. For example, once it is established that there is a red, spherical object in x's field of vision then, given normality of light conditions and of x's sensory equipment, nothing further is required to support the attribution to x of an experience as of there being a red, spherical object before him.

Could a man be in the position of a radical interpreter in

relation to his own experience? A positive answer is available only if it is admitted that a man could have an experience as of an F without knowing that that is the kind of experience he is having. This possibility will be resisted both by those who prefer a doxastic analysis of perception (eg. Howell (1972) pp. 409f) and by those who find persuasive the idea that our beliefs about our experiences are incorrigible. Both ideas are threatened, however, by the lack of evident absurdity in claims of the form : 'x is having an experience as of an F without knowing what an F is', or : 'There is a y which x sees as F but without believing that y is F'. (The quantification into 'x believes that ... ' can be eliminated by whatever method is generally preferred without altering the point.) In principle, then, I might be in the following predicament : I have an experience of an object which is F and this experience is an experience as of something which is F but I do not know either of these facts. I cannot obtain information about the one without obtaining information about the other. In finding out, for example, that the object I am seeing is mauve rather than violet, I will also find out that I am having an experience as of a mauve rather than as of a violet object. The interpretation and the explanation of the experience are delivered simultaneously.

Phenomenalism is reached by arguing that this, or anything else, justifies the claim that there is an entailment from interpretation to explanation. Such a claim is implicit in some arguments against the view that the perception of physical objects requires an inferential process. The mistake lies in supposing that by denying that the interpretative element in perception is a separate cognitive act one has done away with the transcendence of material things. This is a mistake just because, whether or not the interpretation of sensory material is separable from perception, it does not supply a means of escaping from a purely phenomenal space. It is probably true that the initial data of perception are not mere colours and shapes on which we consciously impose a conceptual organisation. But the fact that we unreflectively have experiences as of chairs, rabbits and houses does not itself warrant the assertion that we see chairs, rabbits and houses without having recourse to some kind of induction. The latter claim may also be true, but it depends for its justification on different considerations.

4.3 Once the question of the inferential contribution to the perception of physical objects and their elementary properties has

been settled, a parallel question is raised about the perception of more complex properties, including those of expressive behaviour. Many philosophers, with whom I agree, believe that the recognition of behaviour as expressive is perceptual rather than inferential, that we see anger in a scowl rather than first seeing the scowl and then inferring that the person scowling is angry. Donagan pointed out the problem that this poses for anyone who is concerned to avoid a reductive account of statements about other people's emotions and sensations (Donagan (1966) p. 325). The difficulty is discerned in the context of a discussion of Wittgenstein. But reductionism is not a result of the non-inferential nature of the apprehension of pain in others. It is a result of this in conjunction with a failure to maintain a clear perspective on the distinction between interpretation and explanation, both as it applies to perception and as it applies to descriptions of behaviour. Confusion here closely parallels confusion between the acquisition and the justification of knowledge. Knowledge of others' mental states is obtained, if at all, by seeing their emotions, etc. manifested in their behaviour. We see behaviour as angry, sad, happy. But there remains the question of how this justifies the ascription to people of anger, sadness and joy. The challenge of solipsism, if earnestly pushed, cannot be met at all. It is certainly not met by pointing out that we unthinkingly have experiences as of human beings and not as of senseless bodies about which we conjecture that they may be human (cf. Cooke (1969)).

Behaviour is interpreted when it is described as an expression of something. This may be a description of what such behaviour is typically seen as. Behaviour is explained when its psychological causation is given. This may also be a description of what is seen when sincere expressive behaviour is seen - but it is likely to require a theory of indirect perception which is not required for interpretative seeing. In Chapter 5 I attempt to work out an acceptable theory of both aspects of the perception of behaviour. The effect of this chapter is to stress their non-equivalence. Seeing anger in someone's behaviour is not the same as seeing behaviour as angry. It is theoretically possible that what we see as human beings are merely senseless bodies. It is also possible to see a human being, but to see him merely as a senseless body. A form of behaviourism can be cast as the claim that there is nothing more to an emotion than the physiognomic properties an action is observed to possess.

III. THE EFFICACY OF EXPRESSIVE BEHAVIOUR

1. Behaviour and bodily movements

1.1 The act of smiling is not the same as the smile which appears on the agent's face whenever such an action is performed. The action of gnashing one's teeth is not the same as the resulting gnashing of one's teeth (cf p. 41). The existence of two components in expressive behaviour has been noted by Alston ((1965) p. 18), Ziff ((1958) p. 157 in (1966)) and Wollheim ((1966) p. 93 in (1973)) among others. The latter presents the difference in terms of a distinction between 'activity' and 'trace'; a distinction whose most graphic application is to the plastic arts. Painting is quite naturally understood to comprise two elements; the manual activity which consists in spreading paint on a canvas, and the traces of that activity, the pigmented pattern which remains on the canvas. Similarly, smiling is an activity the product of which is a certain 'lie of the face'. Wollheim stresses that the comparison is not absurd since the difference between physiognomic and artistic phenomena is due to the relative narrowness of the gap between activity and trace in the former, not to the absence of such a gap ((ibid.) p. 97).¹

Designating the two components of expression in this way allows certain interesting questions to be raised concerning the primary bearer of the property of expressiveness and the direction of its 'transmission' between activity and trace. It seems that in artistic expression, the trace is what is primarily expressive, the creative activity being itself expressive, if at all, only insofar as it results in the existence of the expressive trace (Wollheim (1964) p. 281). The opposite is the case with behavioural expression. Here it is the activity which is primarily expressive, while its traces, if it has any, are expressive only in a derivative sense, if at all. Wollheim ((ibid.) p. 280) illustrates this by enquiring as to whether a foot-

1. This seems to represent a modification of the view expressed in an earlier article. In 1964, at p. 285, Wollheim wrote of :

'... the extreme intimacy of the connection between an emotional state and the natural process of expressing it : an intimacy we might characterise by saying that there is no gap between the two.'

print, left by a child stamping with rage, is itself expressive. It seems not to be, constituting merely a 'souvenir' of rage, not an expression of it. However, considerable strain is imposed at this point on the analogy between behaviour and art. The source of the strain lies in the fact that in talking of expression outside the aesthetic context, the subject of our discourse appears to consist in actions. We do not mention traces at all ((ibid.) p. 278). But why, in that case, attempt to preserve the analogy as Wollheim does in the later (1966) paper?

1.2 The difficulty arises because we naturally take the concept of a trace to be the concept of a continuant. But, while in some varieties of behavioural expression certain very transitory continuants seem to be relevantly present, for example, facial configurations, in others there appears to be nothing sufficiently enduring to be thought of as a trace. Gnashings of teeth and stampings of feet are constituted by movement alone. They are, as Wollheim observed, events or processes. This, however, does not automatically mean that they are activities. We need, at this point, the distinction between transitive and intransitive verbs, due to Hornsby (1980), introduced at p. 41 above.² Gnashings_T of teeth are distinct from teeth-gnashings_I. The event of the child stamping_T her foot is not the same event as the stamping_I of her foot. These are descriptions of actions as bodily movements_T, which exemplify the same pattern of entailment as the latter with respect to bodily movements_I: $a\phi_T m \rightarrow m\phi_I$ but it is not the case that $m\phi_I \rightarrow a\phi_T m$.³ Of course, many of the verbs which describe expressive behaviour, physiognomically or otherwise (cf. pp. 18f) do not readily fit this pattern. Consider, for example, 'smile'.

2. I am indebted to Mark Sainsbury for the suggestion that Hornsby's distinction might be relevant here.

3. This is Hornsby's notation. 'a' and 'm' stand for individuals such that 'a' is the subject and 'm' the object of the verb, ' ϕ ', as it occurs transitively, and 'm' is the subject of the same verb as it occurs intransitively. Eg. 'The flames boiled the water' entails 'The water boiled', but not vice versa (cf. Hornsby (op.cit.)p. 124). Note that replacements for ' ϕ_T ' are not always typographically identical to their intransitive counterparts, replacements for ' ϕ_I ' (cf. op.cit. p. 125). Thus, for example, the pair *kill/die*, reflects the same pattern of entailments.

'scowl', laugh', which are always intransitive. The same, however, applies to 'walk'. This need not raise a problem, since when a man walks he moves_T his body in a certain way, and his body moves_I in a certain way (Hornsby (ibid. p. 135). Similarly, when a man scowls he moves_T his face in a certain way and his face consequently moves_I correspondingly.

This will allow the analogy between artistic and behavioural expression to be preserved. In the former, the activity is expressive only insofar as the static object or trace it produces is expressive. The traces of behavioural expression do not disappear, however, but are presented as processes or events, not as static objects or states of objects. It can then be admitted that all there is to behavioural expression is movement, but there are two events or processes involved, one of which constitutes the trace, the movement_I. The question about the transmission of expressiveness will then obtain the same answer in the behavioural and the artistic case. A gnashing_T of teeth is expressive only insofar as the teeth-gnashing_I which accompanies it is expressive.

An alternative strategy, however, would be to isolate three components in both kinds of expression, an activity, a process and a trace (or several traces). Thus there is the activity of putting marks onto the canvas, the event which is the appearing of those marks, and the trace which comprises the marks which appear. There is the act of scowling - the bodily movement_T - the appearance or coming into being of the scowl on the face - the bodily movement_I - and the scowl itself which is a more or less static facial configuration. It is difficult, but not impossible to analyse all behavioural expression in this way, including the earlier examples of teeth-gnashing and foot-stamping. The traces will not be identified as foot-prints or other enduring items, but as fleeting states observable at any moment during the performance of the expressive action. Traces of expressive behaviour are identifiable as aspects of it that can be recorded in still photographs. With traces, events and actions thus identified, the direction in which expressive properties are transmitted between the last-mentioned items in behavioural expression will be the reverse of the direction of transmission in artistic expression. The latter will remain as it was earlier conceived, substituting the process of appearing for the action of making appear. The process of appearing is expressive, if at all, only insofar as what thus appears is

expressive. The hypostasised smile is expressive of joy only insofar as it is part of a process of facial expression, or a configuration which has a certain duration. A smile that is too rapidly terminated or too lastingly fixed is transparently insincere.

1.3 The division of expressive behaviour into three elements is, however, cumbersome and, though it serves certain interests, it is largely superfluous. For the rest of this discussion, therefore, I will assume only two components, both of which will usually be conceived of as events. A scowl is treated as the combination of a facial movement_T and a facial movement_I. The latter, of course, is described physiognomically when it is described as a scowl. Purely by extension from this, so is the former, which thus yields the concept of a 'physiognomic verb'. This calls for a slight modification of Wollheim's definition of physiognomic verbs according to which '... to scowl would be "to produce an angry lie of the face"' ((1966) p. 97). I prefer: 'To scowl' is 'to bring about an angry movement_I of the face', for most occasions, but it would be unwise to force a choice. A person's face might be frozen in an expression of terror in such a way that the very rigidity of the configuration contributes to its expressive power. The idea introduced by the last few paragraphs is that expressive actions are distinguishable from expressive movements_I; that the relation between the action of scowling and the appearance of a scowl on the agent's face is comparable to the relation between a shaking_T by someone of his fist and his fist's shaking_I.

These points serve only to indicate that expressive behaviour conforms, up to a point, to certain familiar theses about action in general. Davidson has shown that all actions are actions of moving our bodies (Davidson (1971) p. 23). This is certainly true of expression. Much of the credit for the idea that many descriptions of actions as anything other than bodily movements are descriptions of bodily movements in terms of their effects is also due to Davidson ((*ibid.*) p. 25). To this, Hornsby has added the observation that even when we describe actions as bodily movements_T, we are describing them in terms of their effects, namely, bodily movements_I ((*op. cit.*) Chapter 1 §4). I have suggested that the same is true of physiognomic movements. More precisely, physiognomic or bodily movements_T cause physiognomic or bodily movements_I and descriptions of events as physiognomic or bodily movements_T entail the truth of corresponding descriptions of other events as physiognomic or bodily movements_I.

There is thus a relation between events and a relation between descriptions of events which together give Hornsby's concept of the 'introduction' of effects by descriptions true of particular actions.⁴ In general a true interpreting description of a particular action, ie. a description of it as expressive, introduces an event of something being expressed.

It is crucial to maintain a clear grasp on distinctions familiar from the philosophy of action when talking about expressive behaviour. The following passage taken from Alston (1965) illustrates the risks incurred by failure to take these distinctions seriously :

The presence of a certain facial expression or a certain demeanour is not a sufficient ground for saying that the *person* expressed determination or indignation, while having said something of an appropriate sort would be. (Alston (1965) p. 18)

The failure of the intended contrast between linguistic and non-linguistic expression is a consequence of the incomparability of the contrasted items. The first claim can be cast in the idiom of the present discussion thus : the observation that a certain expressive facial movement_I has occurred does not provide a (logically) sufficient ground for saying that the person expressed something. The second claim can be rendered thus : if a particular action is of one kind (for example, an utterance-kind) then it is also an action of some other kind (such as an expression of determination). The ground for the latter claim is that all actions of uttering, eg. the sentence 'I am determined', are actions of expressing determination. I agree with this. I also agree that not all determination-expressing facial movements_I are *caused by* determination-expressing facial movements_T and also that no instances of the former are identical to any instances of the latter. But this does not support a contrast between language and expressive behaviour since an emission of sound from a person's

4. In fact Hornsby (p. 70) initially defines introduction as a relation between a description, *d*, possibly true of an action and a further description, *d'*, which must be true if *d* is true. *d* introduces *d'* iff *d* entails *d'*. I do not follow this usage, but reserve the term 'introduction' for the more interesting concept of a relation between a description *d*, true of a particular action, *a*, and some effect, *e*, of *a*; or, as Hornsby puts it, between the pair $\langle a, d \rangle$ and *e*. For the initial concept I will adhere to the concept of entailment. Thus, given a further description, *d'*, true of *e*, we have the definition : $\langle a, d \rangle$ introduces *e* iff *a* causes *e* & *d* entails *d'*.

mouth, even if the sound emitted constitutes a well-formed sentence, need not be caused by a vocal action (or even a speech-act) of which the source of the sound is the agent.

My agreement with Alston extends no further. For Alston needs to deny that the presence of a certain facial expression is a pretty good reason for supposing that the person expresses what his face expresses. The idea seems to be that facial movements_I are not typically, though perhaps they sometimes are, caused by facial movements_T. Emissions of sentence-constituting sounds, however, usually are caused by utterances. The plausibility of the belief that this establishes an opposition between linguistic and non-linguistic expressive behaviour depends partly on the assumption that any sentence which contains a transitive verb and in which the subject-place is taken by the name of a person attributes an action to the person named. The other source of its plausibility is the conviction that actions are properly defined as events which are intentional (under some description). I deny, however, that all movements_T are intentional. Thus it may be that most speech is intentional and that much expressive behaviour of other kinds is not. If this is Alston's point, then, while not fully agreeing with it, I will postpone detailed discussion of it (see Chapter 4 & Appendix). This requires, if only for ease of formulation, a general term for movements_T which are not intentional. Terms such as 'behaviour' are most appropriate if one is to avoid distinguishing a weak from a strong sense of 'action'. In the present context, however, there is little risk of confusion in the adoption of the latter strategy, most of what remains to be said in this chapter being applicable equally to action and behaviour which is not fully intentional.

A related point needs to be mentioned here. A strong case can be made for the withdrawal of interpreting descriptions of movements_I - descriptions of them as expressive - if the discovery is made that they are not caused by movements_T. Part of Alston's point may have been that this is so for language but not for other kinds of expression. 5

5. cf. Kenny ((1976) p. 44) : 'Words written by my hand do not express my desires if my hand was in the irresistible grip of another; if Cardinal Mindzenty at his trial was merely uttering words like a ventriloquist's doll (ie. not himself uttering words at all) then they did not express *his* remorse for his alleged crimes against the state.' The gloss in parenthesis is my own. Note the stressed possessive pronoun; perhaps Kenny would allow that the Cardinal's words expressed remorse, but not remorse actually felt by anyone.

I see no ground for such a distinction, but either way the concept of an expressive movement_T is required if the question is to be raised at all. This question has no bearing on the direction in which expressive properties are transmitted. The suggestion is that a movement_I is expressive only if it is caused by some movement_T, whatever its nature, and not, say, by galvanisation.

2. How basic are interpreting descriptions?

2.1 If perspicuous talk of expressive behaviour is to be achieved then a considerable amount of conceptual material derived from contemporary philosophy of action must be deployed. In particular, the following questions are not only worth asking, but must be kept distinct. These are questions about relations between different items :

(a) The relation between particular actions and particular events which are not themselves actions. If the events in question comprise a bodily movement_T and a bodily movement_I, then the first is the cause of the second.

(b) The relation between action-descriptions and other event-descriptions. For example, 'x moves_T his hand' entails 'x's hand moves_I.'

(c) The relation between descriptions true of a particular action and some effect of that action. If we have descriptions d and d' related as in (b), true of an action, a, and an event, e, related as in (a), then the pair <a,d> introduces e.

(d) The relation between one description of an action and another description of the same action.

Alston failed to distinguish questions about (d) from questions about (a) or (b). There are, indeed, descriptions true of a single action such that one of them entails another. An action, arguably, is a saying that the speaker is determined only if it is an expression of determination (on the assumption that a person will sometimes say 'I am determined' without saying that he is determined).

A more fundamental issue concerns which of two descriptions of a single action is the more basic, and the criteria by which this is determined. The intuitive concept of basicness has its source in sentences of the form : 'x ϕ -ed by ψ -ing'. For example, if I turn on the light by flicking the switch then 'a flicking of the switch is a more basic description of my action than 'a turning on of the light'. The relation ' ... more basic than ... ' is not, of course, a relation

between actions.⁶ Nor is it a straightforward relation between descriptions; a description, d , is more basic than a description, d' , only relative to a single action of which both are true. There could be some other action of which d and d' are also true but relative to which d' is more basic than d (cf. Hornsby (op. cit.) p. 71).

Generalisations about expressive behaviour can be cast in this familiar form. People express happiness by smiling, anger by gnashing their teeth and beliefs by uttering sentences of a language in which they are competent.

A concept of relative basicness has been defined by Davidson on the basis of causation. A description, d , of an action, a , is more basic than a description, d' , of a , iff d describes a as the cause of an event, e , and d' describes a as the cause of an event, e' , and e is the cause of e' (Davidson (1971)). Or again, in the idiom supplied by Hornsby : if d and d' are true descriptions of a , then d is more basic than d' if the effect introduced by $\langle a, d \rangle$ causes the effect introduced by $\langle a, d' \rangle$ (Hornsby (loc cit.) p. 71). The concept thus defined, however, lacks a universal application in the elucidation of 'by'-sentences and should, therefore, be designated 'basic_C' as Hornsby suggests. It seems obvious, *prima facie*, that among the 'by'-sentences to which the concept of basicness_C is inapplicable are included just those generalisations for which an explication is presently sought. The descriptions applicable to a particular action may include 'x's gnashing_T of his teeth' and 'x's expressing his rage', thus introducing the event, the gnashing_I of x's teeth and the event, x's rage being expressed. It is intuitively clear, I think, that x expressed his rage by gnashing_T his teeth; x's action is more basically a gnashing_T of teeth than an expression of anger. But apparently the events introduced are not related as cause to effect. Surely the teeth-gnashing_I and the anger being expressed are the same event. The same applies to every other mode of expressive behaviour. A smile appearing on a face and the happy expression appearing there are identical. The fist shaking_I and the appearance of fury are not separate events. Yet to describe an action as a smile, or a fist-shaking_T is to give a more basic description of it than that given by saying that the action is an expression of happiness or anger.

6. The point is too familiar to require further justification here. It is, in my opinion, conclusively established in Davidson (1971) and Hornsby (1980) Chapter 5, to mention but two possible sources.

2.2 The most natural strategy to adopt at this juncture would be to cast about for an alternative concept of what it is for a description of an action to be more or less basic than some other description of the same action. This would involve the invention of an independent concept or an appeal to some notion already available in the literature not in order to replace 'basic_C' altogether, but simply to accommodate those cases to which it is inapplicable. In favour of this strategy there is the fact that 'by' belongs to a range of quasi-causal vocabulary to which such expressions as 'because' and 'in virtue of' also belong (Hornsby (ibid.) p. 76). To say that these terms are quasi-causal is to say that the concept of causality contributes nothing to the elucidation of many occurrences of them. Some occurrences, of course, are genuinely causal.

Recognising the need for a supplementary concept of basicness if justice is to be done to all the ways in which we discriminate intuitively between kinds of which a single action is an instance, Hornsby supplies the additional notion - 'basic*' ((ibid.) pp. 72-77). This new concept is designed to perform several rather different duties. It must account for descriptions of actions felt to be in some sense their *most* basic descriptions but which are nevertheless less basic_C than other descriptions of the same actions. The intuition that we 'directly' move_T our limbs is well established despite the obvious fact that limb-movements_T are caused by muscle-contractions_I ((ibid.) p. 72). Also, we make finer discriminations between descriptions of actions with respect to basicness than those yielded by the causal concept ((ibid.) p. 74). Finally, 'basic*' is meant to cover suitably modified versions of Goldman's concepts of non-causal 'level-generation': conventional, simple and augmentation generation ((ibid.) p. 75, and Goldman (1970) pp. 26f). It would seem, therefore, that the relation '... more basic* than ...' is amply qualified to explain the non-basicness of physiognomic or interpreting descriptions relative to purely physical descriptions.

However, if our aim is to elucidate the meaning of 'by' in 'He expressed anger by shaking his fist', then the concept of basic* will provide very little assistance. For Hornsby's definition depends on the assumption that 'by' is intuitively clear. Given two descriptions, *d* and *d'*, of a single action, *a*, then *d* is more basic* than *d'* iff *d* appears on the right and *d'* on the left of 'by' in a 'by'-sentence true of *a* (Hornsby (ibid.) p. 73). To say, therefore, that 'his

shaking_T his fist' is a more basic* description than 'his expressing anger' as applied to the same action, is, if true, more or less uninformative. In fact it may be worse than uninformative. 'Basic*' was meant to capture the intuition that an action is more basically described as a limb-movement_T than as a muscle-contraction_T despite the fact that the former is a less basic_C description than the latter. According to a matching intuition, we also want to say that scowling is something we do directly. Suppose we coin the expression 's-movement' as designating in a non-physiognomic way the movements which constitute scowling. Then we do not know how to s-move_T our faces except by scowling. Stuart Hampshire makes the same point by drawing attention to the imitation of facial expression (Hampshire (1960) p. 147 in (1972)). In imitating the facial expressions of others we do not attempt meticulously to reproduce the purely physical movements by which facial expressions are realised. It seems, therefore, that 'x s-moved_T his face' is less basic* than 'x scowled' for much the same reason as 'x contracted_T his muscles' is less basic* than 'x moved_T his limb'. Given the obvious asymmetry of the relation '... more basic than ...' for any one concept of basicness, 'x s-moved_T his face' cannot also be a more basic* description of an action than 'x scowled'.

2.3 The concept of basic* cannot, therefore, be applied in the present context without jeopardising the success of its other applications. A more promising approach might consist in an attempt to modify Goldman's concept of conventional level-generation to meet the present need.⁷ A suitably modified representation of conventional generation will render it as a relation between properties or types of actions, not between tokens. More precisely, an operative convention

7. Goldman ((1970) p. 26) defines conventional generation thus :

Act-token A of agent S conventionally generates act-token A' of agent S only if the performance of A in circumstances C (possibly null), together with a rule R saying that A done in C counts as A', guarantees the performance of A'.

As it stands this definition not only presupposes a theory of action-individuation which is unacceptable to anyone who believes that Davidson, Hornsby et al. are correct on this score, it is also incoherent with respect to individuation given what the relevant convention or rule is supposed to stipulate. For how can an action 'count as' a *different* action from that which it is? 'Counts as' like 'realises', 'constitutes', etc., requires identity of tokens it relates but is asymmetric with respect to the types of which they are tokens.

will be conceived of as ensuring that any action is of a type or kind K, if it is also of some other kind, K'. But the details of such a definition need not be spelled out here because the idea of a convention carries with it the implication of 'psychological interaction' between members of a community. This suggests that it would be a mistake to look for a concept of level generation applicable to the conventional relations between types of which actions are tokens if such a concept is expected to be entirely autonomous from that of basicness_C. Psychological interaction, while not reducible to physical interaction, plainly involves causation. Specifically, to say that actions of one kind are, due to the prevailing conventions, actions of another kind is to say something about what such actions are taken to be by members of the community in which the relevant conventions prevail.

Once this is admitted, the point can be expected to apply to non-conventional types of expressive behaviour such as those under consideration. As a step toward an account of how this might work, one might consider the following example intended by Hornsby to demonstrate an application of 'basic*' ((loc cit.) p. 74). If the sentence 'He greeted her by waving to her' is true of some action, then 'his waving to her' is, according to Hornsby, a more basic* description of this action than 'his greeting her'. But the predicate, '... greets y', could with some plausibility be construed as being applicable only subject to the fulfilment of certain success conditions. He could not be said to have greeted her unless she had been made aware of having been greeted. The appearance of circularity could be dispelled by an approximately Gricean expansion of the notion of greeting thus making explicit the fact that greeting is a form, perhaps rudimentary, of communication. x greets y only if x does something which causes y to recognise x's intention to draw y's attention to x's presence. In Hornsby's example this effect is achieved by waving. Causation is now conspicuously present in the account of greeting and the concept of basicness* consequently becomes superfluous. Her being waved to, which is the event introduced by the description of his action as a waving to her, causes her recognition of his intention as just described, which is the event introduced by the description of his action as a greeting of her. It might be maintained that strictly there are three descriptions of the same action being deployed here. There is one action described respectively

as a 'waving to her', 'a greeting of her' and 'a causing her to recognise that she has been greeted'. But since applicability of the second entails applicability of the third description there is no point in insisting that only the third and not the second description, when paired with the action, introduces the relevant effect.

Still, fists can be shaken and anger thereby expressed in the absence of any witnesses. No effect (other than a fist-shaking_I) is introduced by a description of an action described as an expression of anger; this is hardly surprising, since not all expression is communication. Nevertheless, there is intuitive support for the idea that physiognomic or interpretative redescriptions are comparable to causal redescriptions of actions. I will argue in the next chapter that an action is expressive only if it would produce a certain effect on an observer of a certain kind if one were present. Interpretative redescriptions are a species of conditional causal redescription. But before proceeding to a more detailed account of the conditions under which expressive behaviour is efficacious - to an account, that is, of the relation between the concepts of expression and communication - certain obstacles need to be overcome. Failure to do so would encourage the idea that expressive behaviour is distinguished not by its effects but by its inefficacy.

3. Causation and expression

3.1 Consider the following example. Henry II, in a rage, uttered the words 'Will no-one rid me of this upstart priest?'. Hearing this, four of Henry's knights rode to Canterbury and killed Thomas Becket. There seems to be an unbroken causal chain leading from Henry's utterance - which constitutes, inter alia, an expression of a desire that Becket should cease to exist - to the death of Becket via certain perceptions, beliefs and desires of the knights. The denial of this would be tantamount to a denial that the words of influential personages can alter the course of history. But admission of causal chains of this kind carries certain equally counter-intuitive implications if three general philosophical doctrines are also accepted; doctrines which otherwise have much to recommend them. They are the following :

(i) There are causal links between physical and mental events. (This assertion is intended to be neutral as between theories of psycho-physical causation and theories which depend on the claim that for any mental event there is some neural event with which it is identical.) In this case, Henry's behaviour caused some perceptual

- specifically auditory - event which caused a belief about Henry's desires and complementary desires in the knights which caused their actions which caused Becket's death. Clearly the sequence should be further expanded to include details of the links between audition and belief, but this plays no important role at his point (but is relevant later).

(ii) Transitivity of causation. Applied to the present instance this principle entails that Henry's utterance caused Becket's death.

(iii) Transitive verbs of the kind instantiated by 'to kill' are analysable in terms of causation and their intransitive counterparts. The suggestion is that ' $x \phi\text{-s}_T y$ ' is true iff ' x causes a $\phi\text{-ing}_I$ by y '. To kill someone just is to cause someone to die. Given (i), (ii) and (iii), Henry killed Becket and this is absurd.

One of the above theses must be dropped, or at least, modified if the repugnant conclusion just reached is to be avoided. The most obvious candidate seems to be (iii) since as it stands it is independently capable of supporting inferences of the kind exemplified. The difference between murder and incitement to murder, though perhaps negligible from a moral point of view, is intuitively and legally hard to relinquish. Hornsby has shown how the causal analysis can be modified in a way that will allow the general distinction between $\phi\text{-ing}_T$ and getting someone else to ϕ_T to be preserved. The analysis is retained but conceded to be insufficient. An additional clause is required to ensure that an action is a killing only if the causal route from it to the victim's death does not pass through any other action performed either by the original agent or by someone else. Informally, her proposal is to define transitive verbs as follows : $x \phi\text{-s}_T y$ iff x does something which causes y to ϕ_I and any other action which causes y to ϕ_I also causes x to $\phi_T y$.⁸ This is sufficient to block the inference that Henry killed Becket since among the actions which caused Becket to die is included the knight's action of stabbing Becket which did not cause, though it was caused by, Henry's speech-act.

8. Hornsby ((op.cit.) pp. 124-128) Hornsby follows Davidson (1967) in treating verbs which report actions as two- or three-place predicates for intransitive and transitive verbs respectively. In the first case an event is related to an agent, eg. 'walks (x,e)'; in the second an event is related to an agent and a patient, eg. 'kills (x,y,e)'. Her formal definition of transitive verbs (p. 128) is formulated accordingly, but the details need not detain us here.

There are minor grounds for dissatisfaction here. First, it might be objected that Hornsby's refurbished analysis does nothing to explain why we do not say that Henry killed Becket. It merely repeats that this is so. But then an explanation is not what is aimed for here so much as a semantic description of a certain class of verbs. All that is required is that it should be a faithful record of the way in which the verbs to which it applies actually behave.

Secondly, however, the condition now imposed on the application of transitive verbs may be too stringent. It is sometimes said that a smile launched a thousand ships. Hitler is accused of killing six million European Jews. Yet Helen of Troy never set foot on any of the ships of Argos until after the sacking of Troy and Hitler probably never laid a hand on a single non-Aryan. I suggest, however, that without diminishing the responsibility of Helen or Hitler the sentences 'Helen of Troy launched a thousand ships (by smiling)' and 'Hitler killed six million Jews (by approving the institution of death-camps)' may safely be dismissed as hyperbole. Being less than literally true they do not amount to serious counter-examples to the modified causal analysis of transitive verbs under consideration.

3.2 The causal account of transitive verbs, suitably modified, does not entail that expressive actions, speech-acts and so on, which cause ϕ -ings_I amount to ϕ -ings_T. So far, then, the theory of expression does not require the abandonment of any of the three doctrines mentioned above, but only the improvement of the third. This is a useful result if only because questioning the principle of transitivity of causation would take us too far afield, probably unprofitably. But consider, now, a more mundane example. I ask or enjoin someone to close the door. The person to whom the injunction is addressed complies with the desire expressed in it and closes the door. By expressing a desire for a state of affairs I have caused its realisation. In the light of the previous paragraph this does not, of course, entail that I myself have closed the door. But there remains a certain reluctance to say even that I have caused the door to be closed (by causing someone else to close it). There is, then, a different motive for denying the application of (i) or (ii) above to expressive behaviour. I still do not propose to question (ii), however, because of its general explanatory use.

That leaves the option of denying the efficacy of expressive behaviour altogether. The following passage due to Stuart Hampshire

apparently supports this line.

A scowl, or an angry glance, are allowed to survive, when the rest of the behaviour is inhibited, partly because they are largely ineffective as action. Generally speaking, we effectively do things, and make changes in the world, in the primary sense of these words, that is, in the sense that is associated with physical change, with our hands and with other limbs rather than with our face. When a disposition to behave in a certain way is controlled, the last vestige of the behaviour is apt to survive in the facial expression, and particularly the eyes, as being the ineffective part of the behaviour, the most subtle and insubstantial, and therefore the most immediately expressive of the inner movement of the mind. ... If a movement is seen effectively to serve some evident and familiar human need or purpose, its significance as gesture is lost. The behaviour generally needs to be taken as a sign. (Hampshire (1960) p. 153 in (1972))

When Hampshire says that facial expression is not a way of doing things 'in the primary sense of these words', it may be that he is making, in a very general way, the point made in the previous paragraphs. We do not strike someone by expressing anger and the desire to do so. Henry did not kill Becket by expressing a desire for his demise. But while this much is acceptable, the further apparent implication that it is essential to the notion of a desire ('disposition' or 'inclination') being expressed that the movement by which this is achieved does not cause the desired effect goes rather further. I am not sure whether the implication is intended or is only the result of a misleading way of formulating the generalisation.

This further implication, however, is not obviously wrong. It sounds odd to say that my requesting the door to be shut actually caused the door to be shut. But the oddness pertains to the saying of it rather than to what is said to be the case. It may be true despite its peculiarity. The issue is related to the conditions which determine which of many causal antecedents we pick out as *the* cause of an event. It has been said that the cause of an event is determined in part by the nature of our interest in explaining it.⁹ When several actions cause an event we are ordinarily interested in the action which is temporally last. The notion can be conveyed by noting that to say of an action that it is the cause of an event is usually, *conversationally* to imply (in the now familiar sense) that no other action caused the event and did not cause the first-mentioned action. In other words,

9. cf. Hornsby ((op. cit.) p. 129). The idea is due to Hart and Honoré (1959) *Causation in the Law* (Oxford : Clarendon).

what is conversationally implied corresponds to the additional clause in Hornsby's analysis of causative transitive verbs. What explain the implicature in such cases are the interests presumed to motivate the speaker. The same is true of Grice's well-known example, where a professor conversationally implies that a student's work is inadequate by commenting on the beauty of his hand-writing (cf. Grice (1961) pp. 121-52). The implicature is noticed because professors are expected to be interested in the standard of the ideas expressed by their students, not in the calligraphy of their expression.

Hornsby ((loc cit.) p. 129) has noted the vagueness inherent in the notion of relativity to interest. I would add that it is this vagueness which makes it appropriate in the present connection. Explanatory interests vary, and so, accordingly, do our intuitions about what to count as the cause of an event. Thus, while it is odd to speak of a speech-act causing an event such as the shutting of a door, it is conspicuously less peculiar to say that Henry II's utterance was the cause of Becket's death. Yet there is no relevant logical difference between the two cases.

3.3 The peculiarity of saying that an expressive action, a gesture or a speech-act is the cause of some event caused by an action incited by that speech-act or gesture is thus explained in terms of conversational implicature and interest-relativity. This is preferable to a truth-functional account - ie. an account that would disguise the peculiarity as falsehood - insofar as a proponent of such an account would have either to deny the transitivity of causation or to maintain that what appear in these instances to be straightforward causal sequences are not genuinely causal. The latter option would involve denying that one or more than one of the links in these apparent causal chains is causal. This could be justified only on the basis of some general theory which, for example, denied that desires and beliefs cause actions. In the initial example it would then be argued that the alleged causal route to the killing of Becket by the four knights traverses their beliefs and desires and is not, therefore, genuinely causal. Arguments have been put forward in favour of general theories of this kind, but they are considerably weakened by two kinds of counter-argument. First, if an analysis of causality in terms of nomologically independent counterfactuals is possible then it is futile to maintain in this connection that there are no laws linking desires and beliefs to action (cf. Lewis (1973) pp. 187-8). Secondly, even if this proposal is rejected, the fact that there are no law-like

generalisations linking mental to physical events qua mental and physical events is, even if true¹⁰, hardly sufficient to establish that they are not causally related. For there may be some kind of which a mental event is an instance, instances of which generally cause physical events of some kind instantiated by the relevant physical event. This contention has been argued for most persuasively in recent literature.¹¹

This idea does not mean, however, that the passage quoted from Hampshire's paper contains no more truth than that given in my interpretation of the distinction between expressing and doing things 'in the primary sense of those words'. If this is taken to imply that 'ordinary event causality' is not the principle which determines the route from facial expression to some physical change in the world then it is misleading. Each event in the causal chains under consideration is an instance of some physical kind and is linked in the ordinary way to events which precede and follow it. But if descriptions of what goes on are restricted to this level it will be impossible to capture the way in which such sequences are intelligible as communication. For this an intensional, mentalistic idiom is required and to this extent the kind of causation operating is not 'ordinary'. But its dimension of extraordinariness is supplementary to or supervenient upon ordinary causation. It does not replace it. Expressive behaviour is not, as Sartre (1939) believed, essentially an attempt to alter the world by magical means. It is not ((ibid.) p. 65 in translation) essentially ineffectual.

Hampshire's point is really more subtle. One way of conveying its essence is by distinguishing not between expressing the inclination to ϕ and ϕ -ing but by distinguishing between expressing a desire for some state of affairs which ϕ -ing successfully would realise and trying, perhaps unsuccessfully, to ϕ . This distinction does not necessarily correspond to an observable difference between kinds of behaviour (cf. Gombrich (1972)). For in some cases the specific inefficacy of the behaviour - ie. the fact that it falls short of actually ϕ -ing - is due to circumstances beyond the agent's control. In such

10. Which it may not be - vide Bennett's concept of teleological laws (Bennett (1976) Chapter 2) and their relation to mechanistic laws ((ibid.) §21).

11. For example, by Davidson in (1963) and (1970).

cases, as Gombrich (*ibid.*) puts it, the behaviour will be indeterminate with respect to the distinction between psychological and physical interaction. But while both kinds of behaviour are identically related in one respect to some possible effect - both fall short of achieving it - there is a difference in another respect. The difference is between ineffectually trying to ϕ and ineffectually imitating the action of ϕ -ing. As such, it is plainly teleological rather than causal. More detailed discussion of it is to be found in the Appendix.

4. Conclusion

It is a further corollary of the above and, I think, of Stuart Hampshire's analysis, that the purely physical results to which an action might lead are not those in virtue of which the action is classifiable as expressive. Henry II's utterance would have expressed his wanting Becket to cease to exist even if Becket had outlived Henry. I would have expressed my wanting the door shut even if the door had remained open. Consequently, if physiognomic redescrptions are to be successfully represented as being connected with causal redescrptions, it will not be in terms of purely physical effects, but rather, in terms of cognitive effects. But the relevant kind of cognitive effect has not yet been mentioned. Its specification is the task for the next chapter.

IV. EXPRESSION AND COMMUNICATION

1. Expression as potential communication

1.1 The previous chapter culminated in an argument supporting the claim that expressive behaviour, including speech, could be said to produce concrete effects without logical impropriety and despite the conversational peculiarity of describing facial expressions or speech-acts as the causes of such effects. This was meant as a preliminary to a search for a kind of effect which is such that only behaviour which has the power to produce effects of that kind is expressive. The relevant kind has not yet been specified, but the supposition is that somewhere in the causal chain from Henry II's utterances to Becket's death or from a request for the shutting of a door to its shutting, there is a link which instantiates it. It is reasonably clear that the physical consequences themselves are not what we are looking for. Henry could have expressed his desire for Becket's death in circumstances the same in every relevant respect without Becket dying as a result. (Of course, the circumstances would have to differ with respect to the knights' disposition to do Henry's bidding.) I could have expressed my desire for the closing of the door without the door consequently being shut by my audience. But among the events leading from these expressive actions to the realisation of the desires they express it is likely that there are events such that had they, *ceteris paribus*, not occurred then nothing would have been expressed. The relevant sort of effect will presumably turn out to be cognitive and it is this that underlies the conceptual connection between expression and communication. To exhibit this connection we need first to establish what kind of intuitive support is available. Secondly, an appropriate concept of communication is required, a concept such that :

- (i) any action or behaviour which exemplifies it is expressive,
- (ii) any action or behaviour which is expressive potentially exemplifies it (exemplifies it, that is, under conditions specifiable with sufficient facility to avoid trivialising the connection).

More precisely I will expect of an appropriate concept of communication that whenever 'x expressed F-ness' is true then, if a suitable audience or spectator is present, 'x communicated that Fx' is also true.

1.2 It will be worthwhile to provide some examples which indicate in an approximate way the nature of the intuition expressed in the claim that expression is potential communication. Two introductory examples will suffice.

(i) In the course of an interview a dramatist alleges that his intention in writing a certain play was to express a feeling or vision he had but that he had no intention of communicating anything to anyone. He implies, moreover, that he takes himself to have succeeded in expressing the feeling - angst, let us say - which he set out to express. Before assessing the coherence of the dramatist's claim we need to know what question it was meant to answer. For example, he may have been asked why it is that the play under discussion has only just received its first public performance although it was written twenty years earlier. His answer would then seem to provide a reasonable explanation; there are respectable and familiar ways in which one can imagine it being embroidered. Alternatively, the author may be attempting to meet a certain kind of not unusual, though less than seriously literary or aesthetic complaint. He is accused of having produced a piece which tends to induce disagreeable emotions in its spectators, of having imposed on them a pessimistic vision of life which they would prefer not to acknowledge or even to contemplate. He replies that the work was written because he was afflicted by just such a vision which he felt could be rendered more bearable only through expression in his chosen literary form. He did not, however, intend to induce a similar pessimism in his audiences, though he recognised the probability of this occurring. As long as this apology is not intended as a means of disclaiming responsibility it contains no logical impropriety. One can do something which one believes will produce a certain result without intending to bring about that result but some other which one also believes will follow from one's action.

Finally, suppose that the dramatist is accused of having written a play so utterly obscure that no-one can understand it. It is to this criticism that he replies that he had no intention of communicating anything, only of expressing his own, highly personal anguish about things. He is indifferent to his failure to communicate this feeling, which, he alleges, does not diminish the expressiveness of his work. This contention, it seems, approaches absurdity. If his play cannot elicit some sort of comprehension, then it is a failure not only as communication, but also as expression. In general a purportedly expressive object really is expressive only if there is some spectator

or kind of spectator in whom the object would produce an appropriate cognitive response if it were made accessible to him. This, at any rate, expresses a powerful intuition. Of course, the dramatist is logically entitled to claim that all the actual audiences of the play are composed of individuals too dull to interpret it; that only he, the dramatist, has sufficient wit and erudition to understand his work. But as Wollheim has pointed out ((1964) p. 287) the author's ability to understand his own work must depend on something other than his recollection of how he felt when he wrote it, or even of what he intended to express at that time. This is possible, perhaps, in principle, but a claim of this kind should be treated with suspicion for it is difficult to see how one could verify that the author's interpretation is based on understanding and not on memory.¹ Either way the example suggests that it is a minimum condition of expression that there be some possible audience or spectator to whom it is recognisable as expression.

The example just given supports the supposition of a connection between expression and communication in a literary, ie. aesthetic context. Its bearing on behavioural expression is, therefore, likely to be limited given the differences between artistic and behavioural expression discussed in Chapter 1. The point of the example is to indicate the nature of the conceptual relation between communication and expression at a high level of generality. Once we return to a more detailed discussion it will appear that wherever two concepts of expression differ, as do artistic and behavioural, or behavioural and linguistic expression, there will be two correspondingly different concepts of communication.

Behavioural expression resembles linguistic expression, in one respect at least, more than it resembles artistic expression. Both behavioural and linguistic expression relate what is expressed to the person who expresses it. Artistic expression does not (cf. Chapter 1, §5). The difference concerns the cognitive response produced in spectators. In the aesthetic context that response is essentially contemplative. A work of art which expresses joy invites the spectator to contemplate the idea of joy, not necessarily the artist's joy - perhaps even the spectator's own. Behavioural expression, including speech,

1. The point is strikingly reminiscent of Wittgenstein's critique of the notion of a private language.

is evidential in the sense that it is taken by those who notice it as evidence of some more or less reliable kind about the mental state of the person who exhibits it. Language, in its way, manifests as clearly as art the connection between expression and communication.

(ii) Imagine a love-sick adolescent who murmurs words of passion into his pillow and in the absence of the object of his affection. To say of this performance that it is an expression of love is surely to say that it is behaviour which would be taken by the person to whom it is secretly "addressed" as evidence that she is loved if she were present and if she spoke the same language. It is some such intuition that is expressed, vaguely but unobjectionably, by saying that to express a feeling or belief is to communicate that feeling or belief to a suitably qualified spectator if one is in the vicinity. That this is a genuine intuition rather than a prejudice motivated by a prior commitment to, say, a Gricean theory of meaning is obvious. There is a connection nevertheless. The strength of this intuition demonstrates the weakness of the sort of objections raised against communication-intention theories of meaning which are based on the performance of meaningful linguistic acts in the absence, recognised by the speaker, of any audience (vide Chomsky (1976) pp. 60ff).

1.3 A further highly general point needs to be made about my use of the term 'communication'. I take 'communicate' to be a verb of achievement, two minimum conditions of whose application are the presence of an audience or spectator and the production in the latter of some propositional attitude (cognitive response) vis-a-vis what is said to have been communicated. These requirements are met by any concept of communication essentially applicable to animate entities (as distinct from purely physical concepts such as that of the communication of motion). This is perfectly compatible with the analysis executed by Grice, Strawson, Schiffer et al., of whom it would be misleading to say that they attempt to define communication purely in terms of intentions. What is defined purely in terms of intention is a concept of trying to communicate applicable to a certain range of behaviour which includes speech. Successful communication occurs when the relevant intentions are realised. The concept of expression is the concept of communication with the requirement of an actual audience removed and as a consequence, but only as a consequence of its removal, with the uptake requirement removed.

2. Gricean communication

2.1 Could an account of facial expression be based on a broadly Gricean concept of communication? This suggestion is not entirely unattractive for two initial reasons. First, communication of this kind is not necessarily linguistic, the aim of Grice and his followers being to define a concept of communication primarily and then to analyse linguistic meaning on that basis. Secondly, Grice's aim was to elucidate a concept of meaning - 'non-natural' meaning (Grice (1957)) - suitable for application to language but not only to language. If successful this would support the widely held assumption that non-linguistic behaviour is, or can be, meaningful. A smile means that the person smiling is happy. This idea has exerted an attraction on, among others, Wittgenstein, Wollheim, Charles Taylor and, in a different way, on many anthropologists. The suggestion to be contemplated here is that the meaning of naturally expressive behaviour is 'non-natural'.

The literature of communication-intention theory is replete with examples which seem to meet this desideratum.² An approximate account based on Schiffer ((1972) p. 13) but substituting 'rely' and 'believe' for some of his uses of 'intend' (cf. Bennett (1976) p. 125), can be given for an incomplete analysis as follows :

A person, S, intends to communicate p to an audience, A, by exhibiting a piece of behaviour, α , iff :

- (i) S believes that α has a certain feature, F;
- (ii) S relies on A to recognise that α is F;
- (iii) S intends A to infer at least in part from the fact that $F\alpha$ that (iv);
- (iv) S intends α to produce in A the belief that p;
- (v) S intends A to take A's recognition of S's intention (iv) as a reason for believing that p.

Various explanatory and cautionary qualifications are required here. First I have modified Schiffer's first shot at an analysis of s-meaning to this extent. Schiffer begins by analysing 'S means something by ... ' whereas I have begun with an instance of this general analysandum; 'S means that p ... '. Thus where Schiffer's analysis proceeds in terms of some response, r, the utterer intends to produce via, to coin Bennett's phrase, the Gricean mechanism, mine proceeds in terms of

2. See eg. Schiffer ((1972) pp. 12, 119f) and Bennett ((1976) pp. 13, 138f).

the specific response, a belief that p, which is the proper replacement for r at this stage when dealing with assertion. This is purely for ease of exposition. Secondly, as Strawson, Schiffer and Bennett are all aware, this is not yet sufficient as a definition of s-meaning. More importantly, as Strawson (1964) insists, it is not yet sufficient for the special but not artificial concept of communication that Grice (1957) had in mind. This insufficiency will be relevant, however, only if it transpires that the conditions formulated by Schiffer are necessary for physiognomic and gestural communication. Only then will it be worth attempting here to expand them into a set of conditions sufficient for our purposes. Finally, it is not, of course, communication as such which is determined by the above conditions, but only a certain way of trying to communicate. Communication as such occurs if the speaker's intentions are realised but it should be noted that, for reasons shortly to emerge, this could form the basis only of a sufficient, not of a necessary condition of successful communication. At least the requirement that U's intention (iv) be realised must be dropped and this will have repercussions for the rest of the analysis.

2.2 I have cited Schiffer's modified version of Grice's (1957) proposal because it makes explicit the importance of a certain aspect of s-meaning not mentioned but implicit in Grice. Specifically it stresses that the communicative behaviour must have (or be thought to have) a feature or set of features calculated to fulfil a rationalising role in the production of the audience's belief. This addition is designed to exclude certain kinds of mechanical (eg. neuro-surgical) counter-example (cf. (ibid.) p. 12) which need not preoccupy us here. The relevance of Schiffer's proposal here is that it seems to allow us to sustain both the contrast and the similarity between linguistic and non-linguistic communication. They are similar in that both require complex sets of intentions on the part of the speaker. They differ in that they rely on the audience's recognition of different kinds of property. Thus if S intends to communicate that it is raining to A by uttering 'It is raining', he relies on A's capacity to 'recognise that "it is raining" has the feature of being an English sentence which means "it is raining"' (ibid.). If, on the other hand, S employs the common gesture of repeatedly raising his hand to his mouth in order to communicate to A that he, S, would like to be provided with food, then he relies on A's capacity to recognise that the gesture has the feature of resembling the activity of eating.

Schiffer's suggestion is that :

..... in the first example the relevant feature of S's utterance is a "non-natural" one, whereas in the second example it is a "natural" one. (ibid.)

This promises to supply a basis for distinguishing natural from non-natural or conventional modes of communication, language being an example of the latter and gestures like the one just mentioned providing familiar examples of the former.

The promise, however, will turn out to be illusory. The kind of natural expression with which we are concerned is exemplified by smiling, scowling, fist-shaking, etc.. With respect to the first and second of these at least, there is no obvious 'natural' feature upon which the people who do them rely for the communication of their message. This is an important defect in theories of natural expression which depend upon the iconic or mimetic features of the behaviour thus designated (cf. Hampshire (1960) and also Darwin's 'principle of serviceable associated habits' p. 28 in (1965) edition). Such theories are not easily extended to accommodate the important class of behaviour which, for example, Wittgenstein has in mind when he spoke of natural expression ((1958) eg. §244). As well as the natural expression of pain this class includes smiling as a natural expression of happiness, scowling as a natural expression of anger and so on. The criteria for membership of this class are discussed in the final chapters; they are not relevant here.

Still, while Schiffer's distinction between 'natural' and 'non-natural' features cannot supply a basis for a distinction between natural and conventional expression, his analysis may still be applicable to facial expression. On the model of Schiffer's treatment of communicating something by uttering an appropriate sentence we might say that if S intends to communicate to A that he, S, is happy by smiling then S relies on A's capacity to recognise that smiles have the feature of being a (natural) expression of happiness and he intends thereby to produce in A the belief that S is happy, etc.. Like Schiffer we could assume that the circularity inherent in this proposal will be shown to be harmless by further analysis.

2.3 There is, however, a far more damaging (and not unrelated) objection to the general attempt to elucidate natural expression on the basis of Gricean communication. (Henceforward I will use the term 'natural expression' to denote behaviour like smiling or

or inarticulately crying out in pain. To diminish the risk of confusion I will avoid Grice's and Schiffer's uses of the adjective 'natural' but an analysis of my own use will not be forthcoming until the seventh chapter.) The attempt to formulate a Gricean analysis of facial expression is subject to the criticism that facial expression does not even satisfy the 'sub-Gricean' condition of meaning articulated by Bennett ((*op cit.*) p. 171f). According to Bennett's proposal S means that p by uttering σ only if S intends to communicate p to A by offering him intention-dependent evidence that p. Bennett's aim is to exclude alleged examples of meaning already excluded by Grice and Strawson, which he does on the grounds that they constitute cases where the communicator offers intention-free evidence, while preserving the idea that a community of 'Plain-Talkers' ((*ibid.*) pp. 171f) have a language although they take each other's utterances of σ as evidence that p without recognising speakers' intentions to communicate p. One may not be convinced of the sufficiency of this condition, but its necessity as a condition of s-meaning is beyond dispute. Smiling, however, constitutes evidence that the person who smiles is happy irrespective of whether or not he smiled intentionally. This is so because smiling, like laughter or crying, is often, indeed usually, spontaneous in the sense that the causal route from the emotion expressed (if actually felt) to the behaviour by-passes any formation of intention (communicative or otherwise) by the communicator. Consequently facial expression cannot be elucidated in terms of a Gricean concept of communication.

One way of meeting this objection would be to argue that the contrast it sets up between spontaneous and intentional action is misleading. This is true (and is developed in more detail in the Appendix). and something along these lines must be presupposed by Wollheim, for example, when he claims that to smile is essentially to produce intentionally a certain lie (movement) of the face insofar as it falls under a certain description (Wollheim (1966) p. 97 in (1973)). This is simply to deny that the point of the objection is true. Smiling, like the hand-to-mouth gesture and language, is evidence for the ascription of a state of mind to the person smiling or gesticulating only if it is intentional under the relevant description or, in Schiffer's idiom, only if the person intended to produce a movement which possessed the relevant feature. Bennett's concept of intention-dependence and Schiffer's explicit introduction of features or properties of behaviour into the analysis of meaning can be combined fairly neatly. For a spectator would not infer from the fact that the

hand-to-mouth gesture resembles the activity of eating that the gesticulator is hungry unless he took the gesture to be intentional under the description 'imitation of eating'. This, incidentally, is more than is required by Bennett's sub-Gricean conditions which stipulate only that the evidence is in fact intention-dependent, not that it be taken by the spectator as intention-dependent evidence. This requirement may be too strong for communication which, like speech, depends in part on behavioural regularity. But it seems indispensable for communication which depends on imitation or iconicity.³

The plausibility of Wollheim's position construed as it is here in terms of intention-dependence, is due to the need to exclude from the relevant concept of communication much the same kinds of intuitively non-communicative examples as those that intention-dependence excludes. To arrange convincing-looking evidence that p is not, even if accompanied by the appropriate Gricean intentions, to communicate that p (cf. Strawson (1964) p. 156 in (1971)) unless the evidence is of such a kind that it would not be evidence at all unless accompanied by the right intentions. To offer someone a photograph which confirms the statement that p (cf. Grice (1957) p. 44 in Strawson (ed.) (1967)) is not to communicate that p since the photograph would have retained its value as evidence even if discovered fortuitously. An object or event is intention-dependent evidence for p only if, had its existence been brought about by a purely mechanistic, ie. non-intentional, process, it would not be evidence that p. This requirement seems to be met by smiling as well as by linguistic or iconic communication. A smile produced by galvanisation is not evidence that the person smiling is happy and, given the slightly stronger condition introduced in the last paragraph, if it is known to be thus produced it will not be taken as evidence of happiness.

But it would be wrong to conclude from this that smiling constitutes intention-dependent evidence of happiness. The exclusion of galvanisation reflects only part of what we mean by saying that behaviour is intentional. It is a necessary but insufficient condition of intention that the movements of a person's body should not be caused by the intervention of any external agency. This condition must be met by a facial configuration if it is to constitute evidence

3. Bennett in effect notes the insufficiency of his revised conditions to account for any but regularity-dependent types of meaning ((*ibid.*) p. 174).

for, and hence an expression of a feeling. In short, we require that the smile should be the result of *behaviour*, ie. of a bodily movement_T of the person on whose face it appears. This is born out by the following consideration. There is a consensus among psychologists that infantile smiling is expressive from the age of about four months (see Chapter 7). Earlier apparent facial expressions are merely grimaces mechanistically caused by gastric wind which do not constitute behaviour at all. It would be implausible to suggest that the infant makes the transition from organic mechanism to intentional agency at so early a stage. But it is not obviously implausible to maintain that he begins to behave in certain ways which include a rudimentary form of communication. It is this rudimentary concept of communication which needs to be accommodated in the following analysis.

2.4 The conclusion of this section is negative. It is that physiognomic expression, though it is a species of (potential) communication, is not usefully investigated in the light of s-meaning or intention-dependent evidence (Gricean and sub-Gricean communication). But before looking elsewhere for a suitable concept of communication, a somewhat paradoxical consequence of recent developments in the argument should be made explicit. Bennett's revised definition of s-meaning in terms of intention-dependent evidence, though intended as a weakening of the initial Gricean proposal required in order to accommodate the supposedly linguistic behaviour of unobtrusive plain-talkers, is in another way more stringent than the original account. Bennett's analysis runs as follows : S means that p by uttering σ if he intends thereby to induce in A the belief that p for which the utterance of σ is intention-dependent evidence ((*ibid.*) p. 174 : this is not complete but is all that is required for the present point). This condition cannot be satisfied by smiling. True, it does not follow from this that people never mean anything by smiling given Bennett's strategy of supplying only sufficient conditions for s-meaning ((*ibid.*) p. 22f). But Bennett recognises a need to make his conditions as weak as possible - weak enough, that is, not only to accommodate actual languages, but also communication systems like Plain Talk and non-systematic, iconic modes of communication. It is *prima-facie* reasonable to expect the conditions to accommodate physiognomic and non-iconic, non-conventional gestural communication. But Bennett's proposal makes no sense of the claim that we sometimes mean something by smiling, frowning, raising eyebrows, etc.. For these sometimes constitute evidence for the feelings and attitudes they express without being intentional.

Schiffer's proposal, on the other hand, does admit the possibility of s-meaning through physiognomic behaviour. This would be uninteresting if it amounted to no more than the claim that the elementary phase of Schiffer's theory quoted above could be applied to some cases of smiling. For at that stage none of the counter-examples put forward by Strawson (op cit.), Searle (1965) and Schiffer himself have been eliminated. But Schiffer's strengthening of his conditions, particularly by importing the 'mutual knowledge*' requirement ((ibid.) pp. 30f) does not have the effect of excluding physiognomic expression altogether, though it does achieve its avowed objective. In general, it is reasonable to anticipate that overtly smiling *at* someone in order to communicate some message will satisfy all of Schiffer's requirements, all of them, that is, except the requirement introduced to distinguish linguistic from non-linguistic meaning (see pp. 120ff).⁴ The criterion adopted for the execution of this task is the notion of convention derived from Lewis (1968). Immediately preceding this there is an example of s-meaning minus convention in which members of a certain community utter the sound 'grrr' when they intend to communicate that they are angry relying on the resemblance born by 'grrr' to the sound emitted by dogs when they are angry. The method is mimetic. But suppose that 'grrr' in fact corresponded to a sound spontaneously emitted by members of the same community when they are angry. In that case 'grrr' would constitute intention-free evidence that the growler is angry. But this is no obstacle to their sometimes s-meaning by 'grrr' that they are angry. In the same way people sometimes s-mean something by smiling. Perhaps they often do so if they are civilised adult humans - perhaps, even, with certain kinds of smile they usually do so. This will help to explain the intuition that physiognomic behaviour is genuinely meaningful; meaningful, that is, in a literal, semantic sense. But although, as I have argued, people sometimes, or even standardly, s-mean things by smiling, the possibility of doing so depends on some more primitive connection between smiling and the type of feeling it communicates. A more elementary concept of communication is required.

4. Actually from p. 120 onwards Schiffer is no longer prepared to countenance non-linguistic meaning at all. I will continue to describe cases of Gricean communication which rely on iconic or mimetic features as cases of s-meaning, conforming thereby to Schiffer's usage at pp. 9-10 and Bennett's usage throughout.

3. Communication and belief

3.1 How should we characterise the genus of which what people intend to do when they s-mean is a species? In other words, what is communication in general? Bennett suggests that this general concept of communication is such that 'S communicates p to A' is equivalent to 'S gets A to believe that p' ((op cit.) pp. 127-8). Full Gricean communication is a version of communication in general with added constraints imposed on the method by which the communicative aim is to be achieved. The general concept is intended to be broad enough to cover what Bennett refers to collectively as the 'rejected cases' - notably those which consist in the offering of intention-free evidence that p. While agreeing with Bennett that a broad concept is required, I do not think that he has provided a suitable definition. The condition 'S gets A to believe that p' is supposed to be sufficient and necessary for 'S communicates p to A' in some sense of 'communicate'. I maintain that it is neither.

3.2 First the weakness of Bennett's condition. This is brought out by the following examples.

(i) A neuro-surgeon gets A to believe that p by altering the structure of A's brain. (This is a variant of an example of Schiffer's ((op cit.) p. 12).

(ii) Suppose that p is the proposition that A's house is on fire. A malicious individual, S, gets A to believe that his house is on fire by actually setting it on fire.

It is reasonably obvious that neither (i) nor (ii) are cases of communication in any normal sense. It is also fairly clear how the condition may appropriately be strengthened by means of Schiffer's suggestion introduced at the beginning of the previous section : S communicates p to A iff S gets A to believe that p by offering A evidence for p. This straightforwardly eliminates the counter-examples above while preserving exactly the extension Bennett intended for his general concept. The 'rejected cases', ie. the cases where the proffered evidence is intention-free, all fall within this extension.

A further strengthening of the condition has been suggested by McDowell (1980). McDowell argues that Strawson's (1970) characterisation of a speaker's primary intention (the only intention with which we are still concerned) as the intention to produce in an audience a certain belief is highly unnatural. Its peculiarity is well brought out by its consequences for the conception of interrogation. To ask

a question about a subject is, on this view, to request someone to induce a certain belief about the matter at hand in the interrogator (McDowell (ibid.) p. 126). In order to eliminate this peculiarity McDowell argues that a conception of communication as transmission of information ought to be taken seriously ((ibid.) pp. 125f). To communicate that p is to supply an audience with the information that p . In conformity with this intuitive constraint McDowell suggests that we treat 'communicate' as belonging to a certain class of verbs any member, ϕ , of which class is such that ' $x \phi$ -s that p ' entails the truth of its embedded sentence ((ibid.) p. 126). McDowell claims as an additional advantage of this concept that it reflects the continuity between primitive animal communication systems and human language. If this claim is vindicated then it is just the sort of concept we are looking for since presumably facial expression lies somewhere along that scale. The general notion is as follows : an organism, S , communicates that p to some other organism, A , iff as a result of behaviour exhibited by S , A possesses the information that p . The sense of 'possess information' will depend on the degree of psychological complexity we are prepared to attribute to members of the species to which S and A belong. (In principle they may belong to different species, but I will ignore the possibility of inter-species communication.) For relatively primitive species the concept of information-possession will depend on something like Bennett's ((op cit.) §14) concept of 'registration' along with the truth of the information (misinformation is not a sort of information) and the structure of a causal story relating the subject-matter, the communicative behaviour and the registration. For primitive species again, this story might avoid mentioning evidence - which one would surely want to do for species to which one denies the capacity for belief - by treating their communicative behaviour as surrogate stimuli which elicit the same responses as their referents (cf. McDowell (op cit.) p. 129).

McDowell is aware, however, of the obvious implausibility of applying this concept directly to human, especially linguistic communication. The appropriate concept of information-possession for humans is, of course, knowledge; knowledge that p being analysable in terms of p , belief that p and some sort of causal connection between p and the belief that p . So as far as human communication, at least between consenting adults, is concerned, McDowell's concept of communication is, so far, the same as Bennett's with the additional

requirement that the belief produced in an audience, if one is present, be true. This requirement is obviously not satisfied by every case in which we want to say that S communicates p to A by uttering a sentence in a language of which S and A are competent speakers. McDowell is aware of this and I turn to his response shortly. More directly relevant is the following point: if expression is to be defined as potential communication, that is, as communication with any suitable spectator who may be present, then, given McDowell's analysis so far, 'x expressed F-ness' entails 'Fx'. This was rejected in Chapter 1.

3.3 A weakness of Bennett's condition has been indicated. But it is also far too strong in another sense. The same applies to McDowell's concept insofar as it constitutes a further strengthening. Bennett explicitly represents the production of a belief that p as a necessary condition of communicative success when attempting to communicate that p. This is far too strong since there will be many cases of attempted communication that p, either linguistic or iconic or physiognomic, which do not achieve this goal but which one would normally regard as successful *as* communication. Bennett's definition obliterates the contrast between understanding and believing what is said, a distinction which seems to be applicable beyond the confines of language. This is, of course, none other than a version of the distinction between interpretation and explanation sketched in Chapter 2. To believe what is said in an utterance is to believe that the utterance is explained in a certain way by the speaker's beliefs and that these in turn are explained in a certain way by facts about the world - facts which may, so to speak, be external or internal to the speaker. To understand what is said is not necessarily to have any such beliefs.

Failure to notice this (which is common to Grice, Strawson and Schiffer as well as Bennett) is explained in part by the emphasis on communication-intention rather than on successful communication. The assumption that utterances of indicative sentences are typically accompanied by an intention to induce beliefs in or convey information to an audience is not without plausibility. But it will not provide a basis for s-meaning as long as one allows that people sometimes mean what they say without caring much whether they are believed.

McDowell takes up a suggestion of Searle's (1970) to deal with this. The proposal is to drop Strawson's primary intention and replace it with the secondary intention. The primary intention in speech is to say something to an audience. Communicative success is secured if

the audience is made aware of the speaker's intention 'so that securing mutual awareness is not, as in Strawson's picture, a fallible means to communicative success, but rather constitutes it' (McDowell (op cit.) p. 130). This characterisation of communicative success could still incorporate Bennett's 'rejected cases' and so still meets the requirement he imposes on a general concept of communication. Success in communicating that p is achieved by arranging convincing evidence that p if the audience grasps what it is the communicator is trying to get across. He need not actually be convinced.

One is still entitled to claim the discovery of a concept of communication at once sufficiently integrated and sufficiently broad to make sense of the thought that there is an evolutionary route from primitive communication systems to language; this being one of McDowell's aims. (Whether or not evolution actually followed that route is not a matter for philosophical debate.) In both cases communicative success consists in the transmission of a piece of true (pleonasm) information. In primitive systems the information transmitted refers to the world outside the speaker. In speech it refers to the speaker's intention. In both cases again there is a malfunction or breakdown of communication either if the message transmitted is not received or if it is false; if it is misinformation ((ibid.) p. 134). It might be objected, however, that the comparison between primitive communication and language appears to be worthwhile only given a very vague description of each. True, each essentially involves the transmission of some kind of information, but very different kinds of information are involved in each case. McDowell admits the difference but takes pains to play it down. It is important, however. In a primitive system, if an organism attempts to communicate that p by behaving in a certain way but it is either the case that not p, or that no other organism comes to register that p, then communicative success has not been attained. Not so for language. Primitive communication meets Bennett's condition for communication in general (substituting registration for belief); linguistic communication does not.

I would like to suggest a partial defence of McDowell's claim that his concept is integrated (in a way in which Bennett's is not). The objection rests firmly on the distinction between understanding and believing the content of communicative behaviour. It can be rephrased by saying that in human communication the securing of comprehension is a sufficient condition for communicative success but that in primitive

communication, communication of p is not achieved unless it secures a registration that p in an audience probably of the same species. This is true but misleading. It is true that inducing the belief that p in an audience is not a necessary condition of communicating that p when the method of communication is human language. But it is not true that securing comprehension is not a sufficient condition of successful communication in a primitive system. The point turns on the vacuity of distinguishing belief from understanding in primitive species. For it is difficult to see what non-linguistic behavioural evidence there could be for the assertion that an animal understood the information that p transmitted to it that was not also evidence that the animal believed or registered that p. Consequently, producing understanding of the means of signification is a sufficient condition for communicative success among primitive creatures since there is no empirically discernible difference with respect to such creatures between understanding what is signified and believing it to be true.

3.4 The kind of integrated concept of communication at which McDowell was aiming would be of value in the present context if it allowed physiognomic expression to occupy a transitional position on the scale from primitive to fully linguistic systems of communication. This is a desirable result because of the pretheoretical intuition we have about such behaviour as smiling, snarling, etc.. On the one hand they constitute a kind of natural expression, by which we mean not that their communicative function depends on their possession of "natural" features, but that they are comparable in certain ways to the communicative behaviour of more primitive creatures. On the other hand facial expression seems to be not entirely unlike speech; its expressive properties, and above all the understanding of them, resemble genuinely semantic properties (cf. Wollheim eg. (1964) p. 274). Unfortunately McDowell's concept is not very helpful here since, although it succeeds in representing speech and primitive communication as instances of a single concept of communication, it cannot describe an unbroken transition from one to the other. This is because the options available do not meet both the requirements made explicit in Chapter 1, Section 2 and the requirement specified in Section 2.3 of this chapter. Communicative success in physiognomic expression must not be identified as the successful transmission of first-order information. If it is, then 'x expressed F-ness' will entail 'Fx'.

Nor can we represent the recognition of physiognomic behaviour as expressive in terms of a spectator believing that Fx, perhaps mistakenly; relinquishing, in other words, the condition that the spectator's belief be true. We need the distinction between understanding and believing expressive behaviour. But we cannot either avail ourselves of the notion of the recognition of the communicative intention behind the physiognomy since there may be no such intention. The distinction between understanding and believing must be applicable to the attitudes of adult humans to facial expression, if it is to be applicable at all, to every instance of facial expression they encounter, including that of undeveloped infants. This lacuna in the range of McDowell's concept can be filled, I think, only by incorporating a suitable concept of evidence to be used in the description of the attitude of understanding.

4. Communication and evidence

4.1 A need for the introduction of a concept of evidence into the account of communication in general was recognised at the beginning of the last section. It was found that Bennett's definition was in one way too weak. The notion of communicative behaviour as evidence for what is communicated was brought in to supplement that original suggestion. However, it rapidly transpired that Bennett's condition was far too strong for any worthwhile concept of communication. The next step will be to drop that condition while retaining the proposal initially designed to strengthen it. This gives us : S communicates p to A only if S offers A evidence for p. But the mere availability of evidence is not sufficient for successful communication (though it may be for expression) unless an audience is present who is capable of availing himself of that evidence and does so. Consequently S communicates p to A by means of a piece of behaviour, α , only if A takes α to be evidence that p. This modified conception of communication in general will clearly have ramifications for more specific concepts such as Gricean communication. The suggested modification resembles McDowell's in requiring that we drop the intention taken to be primary by Grice et al. and treat as primary a genuinely communicative intention. Similarly, intention (iv) of Schiffer's definition must be abolished. No doubt this will have repercussions for the rest of Schiffer's analysis but that need not concern us here.

Since the motive for emphasising the evidential nature of communication is to avoid definitions of communicative success with

respect to p which entail either the truth of p or that some audience believe that p , it is obvious that 'evidence' is being used in such a way that : ' p , and p is evidence that q ' does not entail ' q '. In other words, we are talking about prima-facie evidence in precisely the sense elucidated by Peacocke ((1976) p. 172). Prima-facie evidence is evidence which can be overruled without thereby losing its evidential status; in fact it corresponds at least to one ordinary way in which we use the expression 'evidence' tout court. On Peacocke's definition 'p.f. evidence' is primarily a sentence operator which, when applied to pairs of open sentences expresses a relation between 'type-situations'. So Fx may be prima-facie evidence that Gy even if there is some x which is F but no y which is G . But such evidence can be overruled and retain its status as p.f. evidence only if it is overruled by conclusive counter-evidence pertaining to the individual, x , itself. We withdraw the evidential claim altogether if it is found that there is no inferential link between F and G situations in general.

This gives us : S communicates p to A through his behaviour, α , only if A takes α to be prima-facie evidence that p . Before refining this proposal a possible misunderstanding needs to be forestalled. It might be thought that a definition of communication in terms of p.f. evidence could accommodate only regularity-dependent types of communication and not one-off iconic instances. It is easy to see why this is wrong. Any token of the hand-to-mouth type-gesture considered earlier is a way of communicating that the gesticulator wants to be provided with food even if only one token gesture of that type were ever performed. This isolated gesture, if it succeeded as communication, would still do so in virtue of being a token of an action-type which resembles but falls short of eating. In other words, any action that has the property of conspicuously resembling the activity of eating is p.f. evidence that the agent is hungry.

4.2 Let us return at this point to the specific concern of this thesis which is to define what it is for a certain type of behaviour, eg. smiling, to be an expression of a certain type of feeling. Using the concept of p.f. evidence it will be possible to construct a definition in such a way that in the definiendum the phrase ' ... expression of F -ness' will be interchangeable with the phrase ' ... way of communicating F -ness (Fx)' where the definiens consists of some kind of universally quantified conditional. Consider the following

as a first attempt to do justice to the intuitions introduced at the opening of this chapter :

ϕ -ing is an expression of F-ness if and only if $\forall x\forall y :$
 $(x \phi\text{-s} \ \& \ y \text{ notices } x\text{'s } \phi\text{-ing}) \rightarrow y \text{ takes } x\text{'s } \phi\text{-ing as}$
 prima facie evidence that Fx .⁵

As it stands the obvious flaw in this definition is its failure to restrict the domain of quantification. Clearly for any type of expression there will be many individuals who (or, indeed, *which*) are not equipped to interpret it. It might be thought that a solution could be obtained by simply replacing the second universal quantification with the milder requirement that some audience would, if he noticed a ϕ -ing, take it as prima facie evidence of F-ness. This is what seems to be suggested by the examples considered at the beginning of this chapter. However, such a modification of the above biconditional has the absurd consequence that any *unnoticed* action will be expressive of F-ness and that some actions which some observer - eg. a visitor from another planet - takes as prima facie evidence of F-ness would count as expressive on that score alone. A more plausible suggestion would be to introduce a qualification to the effect that only properly equipped observers will, if present, correctly interpret the expressive behaviour. But, while this ought to be a consequence of an adequate definition it is difficult to see how a clause of this kind could be explicitly inserted in the definition without thereby rendering it tendentiously circular. A properly equipped observer just is one who is equipped to interpret ϕ -ings correctly.

These difficulties can be solved by defining expression relative to a group or population. This requirement is obvious enough in the case of linguistic communication and should not meet much resistance when applied to other varieties. Accordingly the concept to be defined becomes : ' ϕ -ing is an expression of F-ness in P', suggesting the following definition :

ϕ -ing is an expression of F-ness in P only if $\forall x\forall y : (x \in P \ \& \ y \in P \ \& \ x \phi\text{-s} \ \& \ y \text{ notices } x\text{'s } \phi\text{-ing}) \rightarrow y \text{ usually takes } x\text{'s } \phi\text{-ing as p.f. evidence that } Fx$.

5. While I attempt to preserve some sort of analogy with Peacocke's use of the concept there is an important difference of aim. Peacocke is not attempting to define any semantic concept but to elucidate what it is for some already interpreted language to be the language of a certain population. I am attempting to define what it is for a piece of behaviour non-interpretatively designated to be the expression of a feeling.

Relativity to a group will be useful later insofar as it will allow the distinction between natural and conventional ways of communicating to be represented in part as based on different extensions of P. Thus ϕ -ing would be a natural expression of F-ness if the condition just given were satisfied when the name of a species rather than a cultural group replaces P. As a limiting case one could allow the possibility of the group containing only one member; the possible identity of x and y. No doubt one should be suspicious of claims which depend too firmly on this possibility, but I am not sure that such suspicion should be manifested in a definition of expression (by stipulating, for example, that $x \neq y$). Such claims are dubious because it would be very difficult to tell of a painter, for example, that his 'interpretation' of his own work was not based solely on his memory of what inspired it (cf. Wollheim (1964) op cit.).

4.3 The things that can be expressed include intentions and desires as well as beliefs and other propositional attitudes. Indeed many emotions essentially involve some libidinal component.⁶ Because of this the condition given above, though necessary, is far from sufficient. It is satisfied by many actions which are not expressive but do constitute and are taken as p.f. evidence that their agents have a certain practical purpose. This could be true of every token of a single action type. For example, whenever a person is seen strangling someone else this will be taken as p.f. evidence that the agent wants his victim dead since strangling generally causes death. Can the suggested condition be strengthened in such a way that it will disqualify ϕ -ing as an expression of the desire to ϕ ? In other words, can we find some way of excluding, in Anscombe's phrase, the 'manifestation' of desires and intentions in the behaviour which they motivate? (cf. Anscombe (1957) §2).

A first point to recall is that from the outset we were trying to formulate a condition on what it is for a kind of physical movement_T - smiling, fist-shaking, etc. - to be an expression of a feeling. This, given the generality of the condition, will exclude the majority of counter-examples of this kind automatically. Most of the things we do are realised by different bodily movements on different occasions. Still, there will remain generalisations of the form ' ϕ -ing is p.f.

6. Some philosophers, including Stuart Hampshire, believe that all emotions involve an inclination to act in a certain way (cf. Hampshire (1960)).

evidence that an agent wants x' which are true just because ϕ -ing generally results in the availability of x . We could, therefore, try to impose an appropriate restriction on the reasons observers have for taking behaviour as p.f. evidence for some kind of mental state. Such a restriction could be derived from the earlier discussion of the inefficacy of expressive behaviour. Indeed we are now in a position to suggest a corollary of Stuart Hampshire's remarks about facial expression (as interpreted in the last chapter). The contrast between primarily effective and primarily expressive or communicative behaviour is that the former provides p.f. evidence about the agent's desires or intentions because of the effects it is seen to have, whereas the latter produces certain effects because it supplies mediating agents with p.f. evidence about the communicator's desires or intentions.

However, this restriction will not fully do justice to the distinction between manifestation and expression put forward by Anscombe. She argues against Wittgenstein ((1953) §647), for example, that stalking is not the natural expression of an intention to pounce among cats. I think she is right about this though not in her conclusion, shared by Stuart Hampshire ((1959) pp. 97-99), that there is no natural, non-linguistic means of expressing intention.⁷ Anscombe and Hampshire would need to support this conclusion by showing why primitive threat behaviour does not express an intention to attack if, for example, the animal's territory is violated or why primitive courtship 'rituals', which are not unlike the speech-act, promising, do not express an intention to mate. The denial that these are natural expressions of intention could be based, as far as I can tell, only on a general apriori reluctance to attribute intentions to languageless creatures.

There are two salient differences between threat and courtship behaviour on the one hand and stalking on the other. First, stalking is a preparatory manoeuvre rather than an expressive gesture. Secondly, stalking would be rendered useless as a predatory technique if it

7. It is odd that Hampshire allows only linguistic expression of intention (and, hence, the attribution of intentions only to language-speakers) given that the account he gives only a year later of non-linguistic expression in terms of the imitation of effective behaviour is most easily and directly applicable to the expression of desires and intentions. To express an intention would be to exhibit behaviour which resembles the kind of behaviour which would, if fully instantiated, realise that intention.

offered its prey an indication that it is being hunted. Stalking is essentially a method of closing with a victim without being noticed. Courtship and threat behaviour, on the other hand, essentially involve display. The difference is fundamentally a difference of function, and this ought to be incorporated into the definition of expression. That is, expression is to be defined not merely as potential communication, but as behaviour whose function is communicative. This can be done without reintroducing communication-intentions given the existence of a non-teleological concept of 'function' (cf. Bennett (op cit.) §23), or, as we might say, a quasi-teleological concept. For, as Bennett points out, it is the similarity of functional to genuinely teleological explanations that partly explain the temptation to describe certain primitive communication systems as languages ((loc cit.) §62). A functional account of a type of behaviour will state that (i) tokens of that type of behaviour produce a certain kind of effect (if performed under suitable conditions) and that (ii) that is why tokens of it are performed.

We want to say that the function of expressive behaviour is to provide members of a group or species with prima-facie evidence for some proposition. The first element in the concept of a function is already present in the condition formulated on p. 84. I suggest that a sufficient condition can be attained by the addition of the following necessary condition :

ϕ -ing is an expression of F-ness in P only if $\forall x : (x \in P \ \& \ x \ \phi\text{-s}) \rightarrow (x \ \phi\text{-s in part because } \forall y : (y \in P \ \& \ y \text{ notices } x\text{'s } \phi\text{-ing}) \rightarrow y \text{ takes } x\text{'s } \phi\text{-ing as p.f. evidence that } Fx.)$

4.4 I doubt whether the counter-examples of the previous paragraphs could have been neutralised by the introduction of a clause about common-knowledge. Anyway, a common-knowledge requirement would be far too strong - indeed there is room for doubt about whether such a requirement must be met by behaviour if it is to count as linguistic (cf. Peacocke (op cit.) p. 174). Language aside, it is certain that ϕ -ing could be a natural expression of F-ness in P even though members of P did not believe that other members of P took their ϕ -ings as prima-facie evidence of F-ness. However it is reasonable to insist that if this were true, communicative success in one sense would never be achieved. Communication involves the achievement of mutual awareness of something between the communicator and the audience (cf. McDowell

(op cit.) p. 130). Given the concept favoured here, communicative success is a matter of mutual awareness not of the communicator's intentions, but that the communicator's behaviour constitutes p.f. evidence that the communicator is in a certain mental state. It should, therefore, be admitted that expression can fall short of successful communication in two ways without losing its title. First an expressive action may be performed in the absence of a suitable audience (member of P). Secondly, an expressive act may fail to secure mutual awareness of its occurrence or nature between communicator and audience, for example because the communicator is unaware that he has expressed anything or unaware of the audience's presence.

5. Some objections

5.1 The point might be raised that in general what gets expressed is not equivalent to what gets communicated. Typically philosophers talk of communicating some "proposition", p, - of communicating that p - whereas we talk of expressing the belief that p. Even if, as I suggested in Chapter 1, we treat 'expresses' as a sentence operator, there will be a difference in most cases between the sentences embedded in 'S expressed p' and 'S communicated q'. The first will always embed a sentence whose subject is S. The second will often embed a sentence which refers to the world external to the communicator. The objection based on this would not merely be that I have assumed the equivalence of p and q but that I have attempted to give a functional account of the relation between expression and communication which this assumption of equivalence renders highly implausible. A plausible functional conception of expressive behaviour in general would surely specify as its primary function the transmission of information about the world, not about the affective condition of communicators since the former makes a more obvious contribution to the survival both of the individual and of the species.

Before suggesting a way of coping with this objection, two clarifications are needed of the contrast between communication and expression on which it is based. First, the criticism clearly does not depend on anything like the old-fashioned distinction between expressive and assertoric uses of language where that is understood as

a distinction between mutually exclusive classes of speech-acts.⁸ The distinction is, on the contrary, between different ways of describing any speech-act. Any utterance of an indicative sentence, σ , which means that p , will be a way of communicating that p and a way of expressing the belief that p .

Secondly, there is no need to insist on the non-equivalence of expressive and communicative content in every case. Avowals are speech-acts in which the utterer both asserts (communicates) and expresses something - the same thing - about his psychological state at the time of utterance. An utterance of the sentence 'I am delighted' is at once an assertion that the speaker is delighted and an expression of delight. We are not forced to say that the speaker expresses his belief that he is delighted. Still, if a high value is placed on the preservation of an invariable relation between expression and communication in every case then one could say this. It is worth noting that Grice's (1968) proposal to define 'S means that p ' as 'S intends to get A to believe that he, S, believes that p ' (along with higher-order intentions, of course) has the effect of equating expressive and communicative content and contrasting them to the assertoric content of all utterances of indicative sentences. Bennett ((op cit.) p 129f) has pointed out that this proposal can cope with avowals only by tolerating the construal of avowals as expressions (communications) of beliefs about the speaker's feelings, but he adds that the repugnance of this consequence is superficial, the incompatibility between belief and incorrigibility being 'merely a matter of suggestion'. The concept of belief only conversationally implies the possibility of error. Bennett further points out that outside the context of fully interpreted languages it is difficult to make out the difference between communicating a feeling and communicating the belief that one has a feeling (ibid.). For the present purpose, then, the two conceptions of expressive content are not only compatible, they are indistinguishable.

The main thrust of this objection, then, is that it is implausible to represent as the *function* of expressive behaviour the transmission of information or evidence about matters internal to the

8. Alston (1965) and Wollheim (1966) have exposed the defects in an extensional conception of the contrast and the distorted comparisons between linguistic and natural expression to which it gave rise.

communicator. There is some truth in this, which could be accommodated in a definition of the expression of feeling by reviving the well-known conception of feelings or emotions like fear, anger, joy, etc. as comprising a doxastic as well as a libidinal component. Fear, roughly, is a sort of belief that something is dangerous and a sort of desire or inclination to distance oneself from the feared object. The primary function of behaviour which expresses fear is to transmit the information not that the communicator is suffering from a certain kind of affective condition but the information that there is something dangerous in its environment. This could be incorporated in the definitions on pp. 84 and 87 by specifying a function, f , relating emotions to types of situation. Thus $f(\text{fear}) = \text{dangerous (type) situation}$. In general, $f(F\text{-ness}) = C$. The definition could then be expanded as follows :

ϕ -ing is an expression of F -ness in P iff there is a type-situation, C , and a function, f , such that $f(F\text{-ness}) = C$ & $(\forall x \forall y : (x \in P \ \& \ y \in P \ \& \ x \ \phi\text{-s} \ \& \ y \text{ notices } x\text{'s } \phi\text{-ing}) \rightarrow y \text{ takes } x\text{'s } \phi\text{-ing} \text{ as p.f. evidence that } C \text{ is instantiated}) \ \& \ (\forall x : (x \in P \ \& \ x \ \phi\text{-s}) \rightarrow (x \ \phi\text{-s} \text{ in part because } \forall y : (y \in P \ \& \ y \text{ notices } x\text{'s } \phi\text{-ing}) \rightarrow y \text{ takes } x\text{'s } \phi\text{-ing} \text{ as p.f. evidence that } C \text{ is instantiated.}))$

In some cases C will in fact be the situation described by ' Fx '. This will be so when some interest (in a quasi-teleological sense) either of the individual or of the species is served by the transmission of just that information - for example the function of expressing hunger by young members of the species is to provide mature and hence more capable members with evidence that they are hungry.

Since in what follows I will be concerned mainly with human facial expression, and since humans, for whatever reason, are supposed to be interested in each other, I will continue to assume that the function of facial expression is to provide others with p.f. evidence about one's feelings. I will treat this as fundamental, though obviously physiognomic behaviour may have other functions in addition to and dependent upon that mentioned.

5.2 A second objection takes issue with my claim that a unified general concept of communication, applicable to primitive and linguistic communication systems, can be formulated in terms of prima-facie evidence. The initial aim was to specify the cognitive response in

terms of the production of which expression can be defined. The relevant response has now been defined as the attitude towards p, '... believes that there is p.f. evidence that p'. This is a sui-generis belief, irreducible to any other propositional attitude. Communicative success is achieved only when the p.f. evidence belief is elicited. But the propositional attitude specified in terms of p.f. evidence is surely too complex to be attributed to primitive, specifically to languageless creatures. Consequently, either I have defined communication in such a way that success can be achieved only with human audiences or I must recognise a different concept of communication applicable to the behaviour of unsophisticated creatures. From the point of view of generality, my account is no better off than McDowell's.

But I offered a partial defence of McDowell which can be modified to meet this objection. True, the belief that certain p.f. evidence supports p is not reducible to the belief that p. It is not even reducible to a disposition to believe that p. Someone who knew in advance that it is not the case that p might still take certain appearances as p.f. evidence that p. But this does not entail that a difference between being disposed to believe and taking something as p.f. evidence that p can always be registered. For cognitively unsophisticated creatures the only behavioural evidence that could support the claim that they take something as p.f. evidence that p would also support the claim that they are disposed to believe (register) that p. And the only evidence in turn that would support the ascription to them of a disposition to register that p would also support (if they are sufficiently unsophisticated) the ascription of an actual registration that p. So, for a cognitively primitive individual there is no difference between its taking p as p.f. evidence that q and its registering that q because it registers that p. The definition of communication in general on the basis of prima-facie evidence does not entail the attribution of complex attitudes to creatures incapable of forming them.

V. PHYSIOGNOMIC PERCEPTION

1. Physiognomic perception and interpretation

1.1 In Chapter 2 I argued that expressive behaviour is behaviour which requires interpretation. Chapter 3 closed with the suggestion that despite first appearances to the contrary, interpreting descriptions of behaviour are not entirely dissimilar from descriptions of actions "in terms of their effects". This suggestion was further elaborated in Chapter 4 where an attempt was made to specify a kind of cognitive response such that failure to produce it, by behaving in a certain way, in a suitably located and qualified observer disqualifies that behaviour as expressive. It seemed that progress towards a definition of the relevant cognitive response could be made by adapting Peacocke's analysis of linguistic understanding in terms of prima-facie evidence. Putting Chapters 3 and 4 together we obtain the idea of communication as a causal sequence running from the communicator's behaviour through the audience/spectator's perception of it to the latter's taking the perceived behaviour as prima-facie evidence that its agent, the communicator, is in a certain affective state. However, there is a further important objection that may be levelled at such a conception of communication and expression. According to the objection I have in mind the foregoing account of expression represents physiognomic understanding as inferential (though not as a complete inference) but such understanding is typically perceptual in nature. Understanding behaviour is not a kind of judgement external to and imposed on the primitive perceptions we have of the behaviour, but is constituted by, or at least firmly based in a kind of perceptual experience peculiar to physiognomy. The claim that there is a special kind of perception involved in the habitual interpretation of expressive behaviour - or, rather, that physiognomic perception is not sharply distinguishable from ordinary perception which is always to some extent interpretative - has a long and distinguished history.¹

1. Some notion of physiognomic perception is defended, stated or merely assumed by the following : Max Scheler (1913), Merleau-Ponty (eg. (1945a) Part 2, Chapter 4; (1945b); (1947); (1960)), Sartre (1943) p. 410, Wittgenstein (1953), Stuart Hampshire (1960), Wollheim (1964; 1966; 1968), Donegan (1966), Cooke (1969), Sircello (1972), Evans and

This would be a sufficient reason by itself for attending to the notion of physiognomic perception, but there are more important reasons. First, it is simply true that, to put it negatively, we do not perceive people's expressive behaviour merely as movements and configurations of varying speeds and durations and then infer from such an impoverished basis a rich set of psychological conclusions. Secondly, a theory of physiognomic perception will have interesting epistemological consequences though, if properly formulated, it will not achieve what has sometimes been claimed for it - namely, an answer to scepticism about other minds. Finally, physiognomic perception, far from being incompatible with the preceding characterisation of expression, fills a hiatus which is indeed present in the causal sequence which I have claimed constitutes communication.

1.2 So far nothing of a positive nature has been said about why we take what we see, when we see expressive behaviour, as prima-facie evidence that its agent feels what is expressed in his behaviour. On the contrary all that has been achieved is the imposition of certain restrictions on the possible routes from observable behaviour to evidential judgements. Facial expression is, or is taken to be prima-facie evidence for what is expressed neither because it is seen to have results which anyone who felt that way would want to obtain nor because it ostensibly resembles behaviour of a kind which would bring about those results. A slightly different explanation which has not been mentioned so far also turns out to be inadequate. Scowling, it might be suggested, is recognised as an expression of anger not because it is or resembles aggressive behaviour - ie. behaviour which produces destructive effects of the kind angry people characteristically desire - but because it frequently accompanies such behaviour. This explanation cannot plausibly be extended to smiling as the expression of, say, contentment or laughter as the expression of mirth. Physiognomic perception, however, indicates that it may be superfluous to look for explanations outside the appearance of the expressive behaviour itself. When people smile they look happy and we take the way things look to be evidence for the way they are.

1. continued

McDowell ((1976) Introduction), McDowell (1978), Aldrich (1978), Charles Taylor (1979). Undoubtedly the list could be considerably extended.

It is clear from this way of putting the matter that a role for physiognomic perception in the understanding of facial expression can be acknowledged without relinquishing the preceding elucidation in terms of *prima-facie* evidence. (In every other respect the remarks made so far leave the concept of physiognomic perception entirely vague.) Physiognomic perception does not constitute understanding but describes the way in which understanding is based in observation. However, this falls short of the claims that sometimes seem to be made for perception. I have given a tentative preliminary account of the relation of understanding as grounded in the perception of the meaning of behaviour. But it often seems to be claimed not only that we see behaviour as meaningful but that we see the emotions themselves in the behaviour.² The two kinds of assertion are rarely sharply distinguished but they plainly differ considerably; they do so along precisely the lines delineated in Chapter 2 where I distinguished interpretation from explanation. In terms of a more or less intuitive ontology, meaning does not exist, being 'primarily a property of behaviour',³ but mental events like beliefs and emotions do. How far is the positive existential claim threatened by the theory of physiognomic perception and the epistemological needs it is designed to meet? The question dominates much of the following discussion.

Distinctions have been made between different senses of 'looks' which may be of assistance here (cf. Jackson (1977) Chapter 2). When a person smiles he, or his face, looks happy in the phenomenal sense of 'looks'⁴ and the physiognomic sense of 'happy' (cf. Chapter 1, §4.1). If I say that someone looks happy to me I do not go beyond a purely phenomenal description of my experience (cf. Chapter 2). Indeed, as Wollheim repeatedly points out, I may be unable to provide

2. References will be forthcoming.

3. Dewey's remark quoted with approval by Quine ((1969) p. 27).

4. This involves extrapolating in a way which Jackson himself might not condone. It is hard to tell because of an apparent inconsistency in his treatment of immediate perception the objects of which (sense-data) may, he allows at p. 22, satisfy predicates other than those of colour, shape and size, eg. predicates which attribute aesthetic or expressive properties. However, though the notion of immediate perception is to be defined partly on the basis of the phenomenal sense of 'looks', Jackson's initial definition of the latter explicitly restricts the descriptions it can contain to those of colour, shape and size (see p. 33).

any other phenomenal description of how he looks which is relevant to the fact that he looks happy (eg. Wollheim (1966) p. 95 in (1973)). But while this may be all that is required for artistic expression, it is crucial to behavioural expression that someone who smiles also looks happy in the epistemic sense; he looks as if he is happy (cf. Jackson (ibid.) p. 30).

Jackson equates 'It looks as if p' with 'certain visually acquired evidence supports the proposition that p'. If this is right then the epistemic 'looks' construction serves to effect a tentative link between the perceptual claims to be scrutinised here and the conception of expression as prima-facie evidence for what is expressed that precedes.⁵ This suggests a very modest interpretation of the claim that we see people's emotions in their behaviour. I see anger in a scowl if and only if the person scowling (is angry and) looks as if he is angry. Whether a more interesting interpretation of the perceptual claim is available remains to be seen. (Whether or not the sentence in parenthesis is required depends on 'grammatical' considerations to be advanced shortly.)

2. Inference versus perception

2.1 The theory of physiognomic perception is characteristically prefaced by a derogation of the role of inference in the ordinary person's unreflective interpretation of his neighbour's expressive behaviour (eg. Donagan (1966) p. 325, Wollheim (1968) p. 81, Evans and McDowell (1976) p. xxii). The basis of this denial is a contrast between perception and inference which contains three specific obscurities (as distinct from its general obscurity alluded to by Wollheim (loc cit.) p. 60), which will bear at least partial illumination at this juncture. First there is no interesting epistemological genus of which perception and inference are contrasting species. Instead a comparison should be made between beliefs which are based on inference and beliefs which are more intimately related to observation. It is not misleading to designate the latter 'perceptual beliefs' as long as the temptation to identify the perceptions themselves with the

5. 'Evidence', as Jackson uses it at this point, means 'prima-facie evidence'. (p. 30)

perceptual beliefs they sometimes deliver is resisted.⁶ A belief, say the belief that Fx , is based on an inference from a premiss, Gy , only if the subject also believes that Gy and that Gy implies Fx . To avoid infinite regress there must be some beliefs which are not formed as the result of an inference. These are the perceptual beliefs, which provide premisses for subsequent inferential beliefs but are not themselves conclusions. Admittedly this leaves the derivation of perceptual beliefs from perceptual experiences shrouded in obscurity, but it does so without detracting from the argument in progress.⁷

The need for a theory of physiognomic perception arises because the habitual interpretation of expressive behaviour is non-inferentially based in the perception of it. Understanding is a kind of perceptual belief and must, therefore, be grounded in a suitable kind of perception. This claim can be supported by transposing certain arguments usually cast in terms of perception itself into the idiom of perceptual beliefs - beliefs about what we see. Wollheim writes as follows :

Indeed it is only a piece of theory, an epistemological presupposition that leads us to think that there is available a neutral description drained of emotion that fits the original perception we have of such objects. ((1966) p. 75 in (1973))

The passage is primarily about the way in which what we see sometimes seems to us to 'match' a certain emotion or mood; Wollheim's proposal is that the perception of expressive behaviour can profitably be compared to the experience of 'correspondences' or 'aspects' in Wittgenstein's terminology. This is suggestive, but the

6. A temptation not resisted by, for example, Howell (1972), who analyses seeing and seeing-as on the basis of seeing-that which in turn he defines in terms of (a modal or possible worlds conception of) belief. Objections to doxastic concepts of perception are to be found in Jackson ((op cit.) Chapters 2 and 7). I have not noticed a tendency explicitly to define physiognomic perception in terms of belief in any of the authors I have cited above. On the contrary, see Wollheim (1980) p. 219.

7. This is a problem for the philosophy of perception in general. I would add, however, that if one accepts my general premiss that for the belief that p to be inferentially based on q the subject must also believe that q and that $q \rightarrow p$, then beliefs can be inferred from perception only if a doxastic analysis of perception itself is true, which I deny in n.6 above. It is irrelevant at this point (though it becomes relevant later) to maintain that perceptual experiences are themselves the bearers of truth-values and thus a species of judgement (Vesey (1956)) or sui-generis propositional attitude (Hintikka (1969) pp. 152-3 and 181 n.5).

epistemological presupposition in question is false only if it is taken to assert that there is always a neutral description of his experience available *to the perceiver*. To make a stronger claim than this would merely be to substitute one unfounded piece of theory for another. Of course, taken literally the rejected presupposition as I have formulated it is rather too obviously false since it makes all perception dependent upon possession of a language; a language, moreover, equipped with an inconceivably rich store of colour, shape and size predicates.⁸ This is born out by the fact that the science of Kinesics, still more or less in its infancy, has had to devise an entirely artificial notation for the description and classification of bodily movements (see Birdwhistell (1971)). But part of Wollheim's point - not, as yet, the specifically phenomenological part but a preliminary to it - can be conveyed in terms of belief. There is no purely morphological perceptual belief which the perception of expressive behaviour delivers and from which we infer that something is being expressed. Consequently, the latter belief must itself be founded in a kind of perception appropriate to it. I will argue later that this does not involve subtracting from, but rather adding to what is traditionally conceived of as the purely "visual" component⁸ of the knowledge or apprehension of behaviour.

Given these qualifications the following statement ought to be taken seriously :

Seeing cheerfulness in a face is not inferring that its owner is cheerful from the way his face looks. (Evans and McDowell (loc cit.))

Of course, interpreted perfectly literally this is uninterestingly true. Vision is not inference in much the same way as thoughts are not sandwich-fillings. That is why I have preferred initially to contrast perceptual beliefs to inferential beliefs. But, to preempt confusion, this is not intended as a denial that certain kinds of visual experience are available only to subjects already in a certain epistemic state derived from an antecedent inference or learning process (cf. Wittgenstein (op cit.) p. 200, Wollheim (1968) p. 77 and

8. Unfortunately the clumsiness of repeated individual references to colour, shape and size is unavoidable in the context of a discussion in which any more general term for these kinds of property will beg important questions. Hence, further down, the scare-quotes around 'visual'.

(1980) p. 221). If this were not the case then Evans' and McDowell's observation would lose its point, their primary interest being the nature of linguistic competence. Similarly, Wittgenstein's analogy between the seeing of aspects and understanding the meaning of a word ((op cit.) p.214) would fail; we learn what words mean and the process by which we do so is comparable in important ways to induction.

2.2 Of the three obscurities alluded to above in the contrast between seeing and inferring, two remain and will be dealt with briefly. First, when it is denied that, as a rule, we infer that emotions are being expressed from morphological premisses about the way people look, this should be understood perfectly literally. In other words, the point is purely psychological and is subject to introspective verification. This might be put by distinguishing inferential processes from inferential procedures. A procedure is understood as a method, a set of rules or, as one might say, a virtual process. As such, a procedure could in principle be realised by a variety of actual processes. For example, a procedure for inferring p from q might be realised non-inferentially in a being in whom the belief that p invariably caused the belief that q through a Humean association of ideas. The process of association is not the same as the process of inference but both might realise the same inferential procedure. Bearing this distinction in mind, if it transpires that there is no reliable procedure for inferring from behavioural premisses to psychological conclusions (deductively or inductively) then we must be content with solipsism. It is no answer to the latter, strictly philosophical position, that we seem to see emotions in behaviour. Phenomenology is irrelevant to scepticism. Nevertheless, it has an epistemological role if we assume that we do know about other minds and then ask how we acquire that knowledge, a question which, since Wittgenstein, has tended to be associated with questions about the acquisition of concepts and the learning of the meaning of psychological terms (see Chapter 6).

Secondly, while it is true that seeing cheerfulness in a face is not inferring that its owner is cheerful, it is also true that the former is no substitute for the latter. This is so not only for reasons advanced in §2.1, but also because inferring that the owner of the face is cheerful is tantamount to explaining the look on his face. But Evans and McDowell were interested primarily in the perception of meanings, ie. the perceptual basis of interpretation

(cf. Chapter 2). We might see cheerfulness in someone's face and believe on the basis of evidence acquired elsewhere that he is miserable. By the same token it would not be superfluous to infer from the cheerful look on his face that he is cheerful. To put the point another way, what is meant by claims that we see something in something else is determined in part by the conclusion of the inferences we are thereby alleged not to execute. Two different claims need to be assessed :

- (i) Knowledge of other people's emotions is perceptual, ie. non-inferential,
- (ii) Knowledge of the meaning of their behaviour is perceptual, ie. non-inferential.

3. 'Seeing-in'

3.1 The construction alluded to in the rubric contains a number of philosophical snares. Ultimately these are best circumvented by abandoning it as a potential analytical instrument applicable to the concept of expression. A useful suggestion of Wollheim's aligns 'seeing-in' with representation and seeing-as with expression ((1970) p. 313 in (1973)). Before following that advice, however, I propose to explain certain unattractive features of theories which appeal to physiognomic perception as being a direct result of equivocations inherent in the 'seeing-in' locution. Its use in the present context has the effect of blurring the distinctions I have urged previously and thus generating the puzzle noted by Donagan in a commentary on Wittgenstein: while it is indeed perverse to describe the "observation" of other people's pain as inferential, the only alternative seems to be that the sentence 'he is in pain' is about observable behaviour (Donagan (1966) p. 325). In short, the notion of physiognomic perception will appear to support a reductive account of the relation between behaviour and mental states if its proponents ignore differences between interpretation and explanation or, again, if it suggests that scepticism can be answered by phenomenology. But the 'seeing-in' construction is too blunt an instrument to record these distinctions with any clarity.

The following is an appropriately illustrative excerpt :

An expression must at least offer what we might call a 'physiognomic reading'. I want to speak of this in cases where we can see X in Y, where X is not identical to Y, and where there is not some other feature of Y, F, which permits us to

infer to X In cases of genuine expression, what is expressed can only be manifest in the expression; whereas in mere physiognomic reading the X we read in Y can be observed on its own. (Charles Taylor (1979) p. 74)

Taylor's schematic letters, particularly 'X', are noncommittal in a number of intriguing ways. Actually 'Y' too is equivocal, but unimportantly so. The Y-position is purely referential and Y can be replaced by individual referring expressions. It does not matter whether the position is occupied by expressions referring to events or to objects; whether, for example, we speak of seeing cheerfulness in a face or cheerfulness in a smile. There is no need to force a decision because to see a bodily movement_I just is to see a part of the body move. There is no perception of events which is not also perception of substances, there being no events without substances irrespective of what we perceive. This is perfectly compatible with the Davidsonian ontology of individual events assumed in Chapter 3. For similar reasons there is no need to worry unduly about whether replacements of Y should refer to bodily movements_I or bodily movements_T, ie. behaviour (cf. Chapter 2 §3.2). Perception of bodily movements_I, of parts of the body moving, constitutes perception of actions when there are actions to be perceived, in the sense that we see actions in virtue of seeing the bodily movements_I to which they give rise (cf. Hornsby (1980) pp. 102f, and for 'seeing in virtue of', cf. Jackson (1977) pp. 15ff and more below).

What urgently requires investigating is 'X', both in relation to its possible replacements and in relation to the position they are to occupy. Taylor's noncommittal notation will be retained at this stage precisely because it allows the following question, among others, to be posed :

(i) What sorts of expression can legitimately replace X?

There are two readily distinguishable alternatives : We can replace X with sentences or with words. We can see 'in another person's facial expression or his behaviour, that he is in pain' (McDowell (1978) p. 136; see also Wollheim (1980) p. 210); we also sometimes say that we can see anger, or his anger, in a scowl. So we can replace X either with sentences attributing emotions like anger to the person whose behaviour we see, or with words like 'anger'. The latter, however, can be interpreted in different ways (cf. Chapter 1). Either 'anger' is a term used to refer to emotions, or it is a nominalisation of an

adjective. And if it is the latter, then it can be understood either physiognomically or in its primary, psychological sense. This gives us four possible interpretations of 'seeing X in Y', three of which correspond, reverting to my usage in Chapter 1, to senses of 'seeing F-ness in Y'. Ignoring for the present the propositional version of the seeing-in construction the difference between nominal and adjectival construals of 'anger' is plainly of the utmost importance, for only the latter supports a physiognomic construal. Thus the seeing-in construction is ambiguous as to whether it applies to the perception of the meaning or the perception of the cause of expressive behaviour: the same ambiguity as that noted in the reservations about inference which motivate the search for a concept of physiognomic perception.

Now it is not part of my contention that one 'parsing' of the X-position is correct or even that at this point a decision as to which is most appropriate can be made; the decision cannot be made in isolation, but only in the context of a series of grammatical and semantic questions still to be posed. But though a decision is not called for, equivocation is still unacceptable and the passage quoted is equivocal almost to the point of incoherence. For since it is a fact about physiognomic reading that, according to Taylor, the X we see in Y can also be seen on its own in cases other than genuine expression, X must be an object of some kind - presumably an event. This is strongly suggested also, but not entailed by the statement that X is not identical to Y. On the other hand the force of 'other' in 'there is not some other feature of Y, F ... ' is inevitably 'other than X', indicating, contrary to the previous suggestion, that X is a feature, ie. a property of Y. The equivocation persists in Taylor's non-schematic illustration of physiognomic reading which is not expression (Taylor (loc cit.)) : we see instability in a building so the illustration has it, without being able to specify those features of the building which make us think it is unstable. But the building does not express its instability because if we continue to watch it we will eventually be in a position to witness its actual collapse. What we see in the building, then, is both its instability, ie. a property of it and its imminent collapse, ie. a future event. It is, of course, all too obvious that there can be (apparent) instability without imminent collapse - witness the Leaning Tower of Pisa, and imminent collapse without instability, as suffered by many buildings

in London during the Blitz. But it is apparently less obvious, at any rate to many philosophers, that there can be expressions of sadness without any sadness; that while meanings, including what is attributed by 'sad' physiognomically construed, are "mere" properties of behaviour, emotions are not. The latter are causes of behaviour.

3.2 Before pursuing this grammatical scrutiny of the claim that we see anger, joy, sadness, etc. in people's expressive behaviour, a short interlude is needed in order to anticipate a likely criticism according to which my interpretation of Taylor's article is uncharitably literal.⁹ Certain points can be made in defense of the admittedly somewhat laborious approach of the previous paragraph. First it would be equally uncharitable to dismiss as metaphor what is plainly intended as a piece of philosophical writing. This applies to all those mentioned on p. 92, and not only to Charles Taylor. Aldrich (1978) has forcibly asserted the literalness of the preposition in the construction in question : in doing so he speaks for most of those who use it. And it is crucial to McDowell's argument against anti-realism with respect to other-ascriptions of psychological states that 'one can *literally* perceive, in another person's facial expression or his behaviour, that he is in pain' (McDowell (1978, op cit.) p. 136). True, some philosophers have attempted to elucidate expression in terms of metaphor (notably Goodman, cf. Chapter 1) and the concept of metaphor has sometimes been exploited for the purpose of understanding physiognomic and other extensions of ordinary perception (eg. by Howell (1972) p. 412). Aldrich and McDowell are probably right ; the introduction of a concept of metaphor at this point has the air of an arbitrary prohibition of further analysis. Such a prohibition could be justified only by a demonstration that analysis beyond the point reached must lead to an impasse, and this has not, as far as I know, been supplied.

However, the burden of the objection may have nothing to do with metaphor, there being other non-literal uses of language. Apart from obvious figures of speech - hyberbole, irony and so on, there may be imprecise idioms which are seriously misleading only when we try to analyse them through decomposition rather than paraphrase. Imprecision of this kind, which has nothing to do with vagueness, is

9. Mark Sainsbury objected on this ground to an earlier draft of this section, much of which has now been suppressed or replaced with new material. I suspect, however, that the objection might still be considered apposite.

acceptable where there is little risk of genuinely confusing items of different kinds such as instability and imminent collapse. But the concepts associated with expression are less immediately transparent and loose talk could cost us our mental lives.

There is a real difficulty here, however, which is not captured by contrasting metaphorical and literal uses of language, or precision and more or less tolerable imprecision, but concerns both the radical difference and the interplay between phenomenological discourse and ordinary objective language.¹⁰ Perhaps it is simply to miss the point to insist on proper use of the apparatus of reference and quantification when what is aimed at is a purely phenomenological description. But pure phenomenology, where existence is put aside in a Husserlian parenthesis, is rarely done without an eye to some concrete philosophical goal. Clearly any hope of attaining such a goal will be dashed should it transpire that the results of phenomenological investigation in general are incapable of communication through the usual philosophical and objective idioms without seeming to generate esoteric entities on the one hand or eliminating respectable ones on the other. The basic concept of perception itself involves the phenomenal and the real, the way things seem and the way they are in various permutations which it is the task of the theory of perception to define. When dealing with the more recondite perceptual idioms relevant to the present concern those permutations may be subject to alteration. For this reason we need to conduct a grammatical investigation, however pedestrian it may seem.

3.3 §3.1 indicated that there are four possible interpretations of 'A sees X in Y' with respect to possible replacements of 'X'. But nothing is settled by a decision in favour of one rather than another until that decision is related to a number of other grammatical and semantic choices. In addition to the first question with respect to 'A sees X in Y' :

(i) What kinds of expressions replace 'X'?

we have the following supplementary questions :

(ii) Does 'A sees X in Y' entail 'S sees X'?

(iii) Is 'sees' here used with its ordinary grammar?

10. A theme which deeply preoccupied Wittgenstein in *The Philosophical Remarks*.

- (iv) How many objects of sight are there?
- (v) How immediately are they seen?
- (vi) What does 'in' mean here?

This gives us six questions in all which need to be approached to some extent holistically in the sense that only series of answers taken together will yield determinate truth conditions for the kind of sentence in question and in the sense that overall permutations with different atomistic components will sometimes yield practically equivalent interpretations of 'A sees X in Y'. It is of course unnecessary to mention every possibility; only those which yield interestingly false or promisingly true statements need be considered.

The most interesting initial hypothesis is that (ia) expressions whose ordinary function is to denote emotions can replace X. We can supplement this with the further hypotheses that (iia) 'A sees X in Y' entails 'A sees X' and (iiaa) 'sees' is being used in the usual, extensional way. In other words 'A sees anger in somebody's scowl' entails that there is somebody's anger which A sees. This constitutes a partial interpretation which can be made to yield a complete definition in only one of two consistent ways. One complete interpretation continues as follows : (iva) there is only one object of sight; (va) which is seen with whatever degree of immediacy characterises the seeing of objects in general which belong to the same ontological category, ie. mainly events; (via) and 'in' serves to relate the seeing of X to the seeing of Y on the basis of the identity of X and Y. (ia) to (via) will provide a definition somewhat as follows :

Di : 'A sees X in Y' is true iff $Y = X$ & A sees Y with
a reasonable degree of immediacy.

This is, of course, incompatible with Taylor's requirement that we refrain from identifying X and Y. The only alternative that retains (ia) to (iiaa) seems to be the following : (iva) there are two objects of sight, X and Y; (va) which are such that X is seen less immediately than Y; (via) and 'in' is equivalent to 'in virtue of seeing', relating the seeing of X to the seeing of Y on the basis of some relation between X and Y other than identity. The concepts of relative immediacy and seeing something in virtue of seeing something else are derived from Jackson (1977, loc cit.); more will be said about them shortly. We now have a second definition :

Dii : 'A sees X in Y' is true iff $r(X,Y)$, $X \neq Y$ &
 A sees X in virtue of seeing Y.

Remaining at a level of abstraction at which the relation, r , is left unspecified, we already obtain an apparent dilemma which resembles in some respects that which perplexed Donagan (loc cit.). According to Dii we perceive other people's emotions, pain and the like only mediately which is apparently what Wittgenstein sought to deny. The only alternative involves identifying emotions with behaviour.¹¹

3.4 There are two reasons for not worrying excessively about this dilemma :

(a) Immediacy of perception is plainly a matter of degree when it is defined on the basis of the 'in virtue of' relation. Our perception of other people's emotions is mediated by perception of their behaviour, which in turn is mediated by perception of their bodily movements_I. The perception of objects and events in general is similarly mediated and, except in relation to a single series of objects corresponding to one perception, there is no reliable method for comparing degrees of immediacy with which different objects are perceived.

(b) To say that an object is perceived only mediately is not to say that its perception is dependent upon an inference. Paraphrasing Jackson's ((op cit.) pp. 15-20) explanation of mediated perception, x is a mediate object of perception if and only if there is some further object, y , in virtue of perceiving which we perceive x . We perceive x in virtue of perceiving y only if there is a suitable relation between x and y ; for example, y is a part of x .¹²

11. Explicit assent to such an identification is rare for obvious intuitive reasons. An apparent exception is to be found in Merleau-Ponty ((1948) pp. 48-62).

12. A's seeing y and y 's being a part of x obviously cannot be sufficient for the truth of 'A sees x in virtue of seeing y '; if it were, then we would see the entire universe whenever we saw an individual object, a conclusion whose Neo-Hegelian extravagance will probably not recommend it to contemporary ways of thinking. What determines decisions as to whether we see what we see a part of may well involve relativity to interest of the kind earlier tolerated in connection with decisions as to whether an action counts as the cause of a subsequent event (cf. Chapter 3, pp. 62-3). The difficulty and probable vagueness here will not influence subsequent developments.

Most physical objects are such that we perceive them in virtue of perceiving their surfaces and we perceive their surfaces in virtue of perceiving parts of their surfaces. Given that Jackson also holds that we perceive material objects in virtue of perceiving sense-data, and the presence of a perceived object causes the appearance of a datum, causation is also a suitable relation between objects if one is to be seen in virtue of another. Less controversial examples might include seeing an object in virtue of seeing its reflection in a mirror. There is, therefore, nothing to be gained in explaining the perception of emotions by claiming that appropriately expressive behaviour is a part of the emotion expressed. (An analysis along these lines is suggested by Tormey ((1971) pp. 48f.)

In the light of this concept of mediated perception it seems plausible to assert that we do perceive other people's emotions. Dii, therefore, translates 'A sees X in Y' as a reasonable statement where it concerns the seeing of emotions as such in relation to the perception of behaviour. Moreover, if we see emotions at all, then we certainly see them only mediately in virtue of seeing behaviour. However, for reasons evidently related to what precedes, Dii is inappropriate as a definition of physiognomic perception. First, just because Jackson's concept of an immediate object of perception is not contrasted to some notion of an inferred object of perception, it is irrelevant to problems of physiognomic perception. The fact that emotions are at best mediate objects of perception does not entail that we infer their existence from the behaviour we perceive. On the other hand, no alternative explanation of how we recognise that we are perceiving emotions has been suggested. Secondly, physiognomic perception is supposed to correlate pretty closely with expression, even if the most that can be claimed is a purely de facto correlation true of human beings (cf. Wittgenstein (1958) p. 218). But there is not an emotion to be seen for every physiognomically perceived piece of expressive behaviour given the possibility of insincere expression. The latter possibility demands an explanation of the way in which sincere and insincere expression look the same which goes beyond purely morphological description (for reasons which will be clarified shortly). Jackson's definition of mediate perception is silent on the issue of how things seem insofar as there need be no match between my visual experience and an object which causes it for it to be true that I perceive the object. Conversely, the theory of mediate perception is

insufficient by itself to accommodate the sense in which I am the victim of an illusion, not merely of erroneous belief, when it seems to me that there is before me a physical object of a certain kind, there being in fact only an imitation of the surface of objects of that kind. Similarly, if emotions themselves are to be represented as being more accessible to observers than the model of inference from premisses about behaviour suggests, then the primary effect of insincere expressive behaviour is to generate the illusion of an emotion.

3.5 The argument of the previous paragraphs could have been based on different initial hypotheses. With certain added complexities we could have examined the claim that we see that someone is angry in his scowl, assuming that 'sees that' occurs in this context with its usual success grammar as we assumed that 'sees' functioned extensionally. The other questions (iv) - (vi) would have to be construed then as questions about whether we are dealing with a simple instance of epistemic seeing, or a case of secondary epistemic seeing in relation to emotions, based on primary epistemic seeing in relation to behaviour (cf. Jackson (op cit.) pp. 159-167). Again, questions about the perception of properties are reducible either to questions about epistemic seeing or to 'seeing-as' which I propose to consider shortly.

4. Physiognomic versus representational seeing

4.1 The expression, 'representational seeing' is borrowed from Wollheim (1968); it is a convenient phrase for designating en bloc the way in which we see not only people and things in representations of them, but also in flames, clouds and ink-blots which lack 'artificiality' and are not intentionally produced as representations. The details of representational seeing lie beyond the purview of the matters at hand but some comment, again negative, is called for in view of the powerful attraction exerted on many by analogies between expression and representation, physiognomic and representational seeing. A paradigm for this tendency finds expression in Wittgenstein's phrase: 'The human body is the best picture of the human soul.' ((1958) p. 178). The analogy persists throughout Wittgenstein's discussion of the semantics of 'pain' and psychological terms generally in terms of 'Bild' and 'Vorstellung' (§300), and receives an interesting refinement at §297 (see also Donagan (1966) for a worthwhile attempt to translate Wittgenstein's more cryptic pronouncements into a more literal idiom). There are, however, grounds for doubting that we see pain in pain-

behaviour in the way in which we might see a person in his portrait. There are grounds, too, for suspecting that this analogy underlies much of the obfuscation in this area.

In terms of the scheme outlined in §3.3, representational seeing obviously requires a very different analysis from those considered hitherto. Still, the initial hypothesis can be retained, being if anything more plausible here than it was earlier : 'X' in 'A sees X in Y' is replaceable by singular referring terms, for example, by proper names in the case of portraits. The only other alternative here is the propositional version (cf. Wollheim (1980) p. 210), but for simplicity I will focus on the 'direct object' construction. Moving on to phases (ii) and (iii) we can say either that (iib), 'A sees X in Y' does not entail 'A sees X' where (iiib), 'sees' is being used extensionally only in the second sentence ; or we can say that (iic) 'A sees X in Y' does entail 'A sees X' but (iiic) 'sees' is not being used in the ordinary extensional way (cf. Ishiguro (1967) p. 44f). Since (iib) and (iiib) either lead to a *cul de sac* or require supplementation by means of (iic) and (iiic), speculation is safely confined to the latter. This does not mean, however, that we cannot eliminate the intensional 'sees' in favour of some construction employing its ordinary extensional counterpart; but such a construction will still need to contain replacements for 'X' non-referentially, eg. by introducing the verb after a reference-cancelling expression like 'It's as if ... '. Proceeding, now, to (iv) a choice is available : either (ivc), there is only one object of sight, which is whatever is denoted by Y, or, as Quine puts it, extensionality can be nominally restored by (ivd) introducing an intensional object as the referent of X (cf. Ishiguro (*loc cit.*) and Wollheim (1980) pp. 213ff). The question of relative immediacy, construed as above, lacks application certainly if there is only one object and, as far as I can tell, also if we allow intensional objects of representation which, for reasons which will become clearer if they are not already obvious, cannot be identified with sense data.¹³ Since intensional objects are now firmly established for better or worse the function of 'in' can be construed accordingly. Aldrich (exploiting an idea of Tormey's) suggests that 'in' is disclosed by 'prepositional analysis' as

13. If they could, then we would see the X we see in Y more immediately than the Y in which we see it; if sense-data exist then there is nothing we see more immediately than sense-data.

expressing the ontological dependence of the intensional object, X, upon Y : this dependence is captured by the Meinongian concept of 'intensional inexistence'. 'X inexists in Y' (Aldrich (1978) p. 214).

4.2 This is, of course, the merest skeleton of an analysis of representational seeing. But it includes sufficient material to be unacceptable to those who, like Goodman, favour austerity at all costs, and to be recognisable as corresponding to elements of theories put forward by Ishiguro, Wollheim and Aldrich among others. Moreover, the type of analysis adumbrated just now needs very little filling out for it to be clear how misleading its application to expression would be. For whatever relation there is between Y, the painting, and X, a unicorn, griffin, Prince of Wales or Churchill, it cannot correspond to the relation either between behaviour and the felt emotions it expresses or to the relation between behaviour and its emotional 'meaning' or physiognomic character. First, there is a difference between, to put it crudely, what is in a painting and what a painting is of. Unicorns and Griffins correspond only to the former. They do not exist outside paintings of them or stories about them. Churchill, on the other hand, was a real individual. As Goodman puts it, to say of a painting that it is a representation of Churchill is to say, first, that it 'denotes' the individual, Churchill (Goodman (1976) pp. 21f). But there are no unicorns to be denoted by representations of unicorns. Goodman suggests that we construe statements about the representation of non-existent objects as specifying not the referent of the painting but a kind to which it belongs (loc cit.). We should reserve the expression 'representation of x' for genuine denotational representation, rewriting sentences about non-denotational representation by means of the expression 'x-representing picture' or, more briefly 'x-picture' (loc cit.). Now most representations of Churchill will turn out to be Churchill-pictures providing ample scope for ambiguity particularly when unanalysed notions of representation occur as analogues in the elucidation of expression. Failure to distinguish between these two senses of 'representation' can lead to ontological disaster. If non-denotational is assimilated to denotational representation, one needs to postulate a referent for pictures of unicorns. If, on the other hand, denotational is assimilated to non-denotational representation, then the represented object disappears or retains only a somewhat diminished reality. If one applies the first assimilation to the relation between, for example, anger and scowls,

one of two conclusions must follow. Either anger is the object which scowls invariably 'denote', in which case there could never be a scowl without anger; or scowls sometimes denote angry emotions and sometimes something more doubtful ontologically. If the second assimilation is adopted, then anger is reduced to a property attributable to certain kinds of behaviour or facial configurations. In general, failure to keep Goodman's distinction in mind has the result that comparing expression to representation will make mental events seem as suspect ontologically as meanings.

4.3 Neither denotational nor non-denotational representation will support a concept of representational seeing applicable to physiognomic perception or the perception of emotion in virtue of the perception of behaviour. This is bound up with the extraordinary difficulty of eliminating the intensionality of representational perception at least by techniques which have some measure of success in application to physiognomic perception. This difficulty is undiminished, or only slightly diminished, by Goodman's extensional treatment of both varieties of representation, treating the one as a matter of reference and the other in terms of membership of a class. The concept of expression is divided in an analogous way, though this is a rather superficial observation (cf. Chapter 1.). At all events, the analogy cannot be extended to the pertinent kinds of perception. There is a temptation here which must be resisted. The seeing of emotions in virtue of seeing appropriate expressive behaviour and the perception of the physiognomic properties of the latter cannot be assimilated respectively to seeing the man, Churchill, in a portrait of him and (non-extensionally) seeing Churchill (or a unicorn) in a painting. For when we see Churchill in a painting we do not see the man, Churchill, at all. One cannot see x in virtue of seeing a representation of x ; this statement ought to be supported by a sufficiently detailed analysis of causation in perception. I do not know whether the causal theory of perception has achieved this but it is clearly an intuitive desideratum. Little scrutiny is required to disclose the absurdity in a statement like the following: 'I saw my father in Hampstead yesterday. I saw him again today in that portrait in the hall'.

The other part of the analogy might seem to be better off. Representational seeing in relation to the non-denotational aspect of representation can be construed on the basis of the seeing of properties - seeing-as. This cannot be a simple matter of replacing seeing x in y

with seeing y as x (cf. Wollheim (1980) p. 226); if it were, then whenever I saw a picture of a unicorn I would be subject to an illusion (as Gombrich thought). The error is sometimes further compounded when it is thought that seeing y as x , or, more naturally, seeing x as an F involves believing that x is an F . Thus Howell (1972) maintains not only that seeing the duck-rabbit as a rabbit is to be deluded, but also that it is to be deceived into believing that the object before one is a rabbit. ('Duck-rabbit', of course, denotes a kind of drawing, not a mythical creature like a griffin.) These are oversimplifications, but there may be a systematic correlation between classes of seeing-as statement and classes of seeing-in statement. Ishiguro, for example, regards the claim that I see a painting as a picture of A as equivalent to the claim that I see A in the picture ((op cit.) p. 46). Combining this with Goodman's concept of non-denotational representation we can equate claims to see unicorns in paintings with claims to see paintings as unicorn-pictures. Analogously the physiognomic perception of expressive behaviour is a matter of seeing behaviour as angry, which is equivalent to seeing it as an expression of anger (Chapter 1.).

However, this strategy works better for physiognomic perception than it does for representational seeing for reasons bound up with what Wollheim ((1980) p. 213) calls the 'two-foldness' of the latter. Representational seeing just is seeing things as representations, but seeing something as a representation requires further decomposition into two elements one of which seems irreducibly a matter of what we see in representations. The attempt to eliminate the intensionality of representational seeing may ultimately be futile. It is not, however, within the competence of this discussion to make a final decision; it is sufficient to indicate that the comparative difficulty of imposing extensionality on representational seeing is not matched in the case of physiognomic perception and that this is so because of a concrete discrepancy between the two concepts. Wollheim's concept of two-foldness is a response in part to the ineptitude of construing the perception even of highly "naturalistic" art in terms of illusion. The same scruple need not apply to expression - at least not to behavioural or facial expression. When I see a unicorn in a unicorn-picture, then not only do I not see the picture as a unicorn, but it does not seem to me as if I (extensionally) see a unicorn. More briefly, it does not seem to me that there is a unicorn before me. But when I see anger in a scowl, then it does seem to me that I

am witnessing someone's anger, that anger exists before my eyes in the only sense in which an emotion can exist; it seems that someone is angry. If the notion of physiognomic perception is to contribute epistemologically to our understanding of expression and communication, and ultimately to the processes of intersubjective cognition, then cases of seeing anger in a scowl where the scowler is not angry must be plausibly classifiable as illusion.

5. Seeing-as and physiognomic illusion

5.1 I turn now to Wollheim's suggestion, alluded to at the opening of §3.1 :

' ... what we see in a picture (partially) determines what it represents and what we see it as (partially) determines what it expresses.' (Wollheim (1970) p. 313 in 1973)

The nature of the contrast between seeing-in and seeing-as is discussed in more detail in Wollheim (1980), but, since to do a proper phenomenology of the distinction would take too long and anyway would not conform to the overall tenor of this essay, I will confine myself to making a few, relatively abstract remarks.¹⁴ The main points about physiognomic properties, ie. what we see expressive behaviour as, are (a) that physiognomic properties are irreducible to any other kind of visually accessible properties; (b) that, nevertheless, they have the same *phenomenological* status as colour, shape, etc.; but (c) they have a different *ontological* status.

These observations deserve some clarification, though whatever justification is available must be allowed to remain implicit in the descriptive use to which they are put. The alleged irreducibility of physiognomic predicates is based on the *prima-facie* hopelessness of any attempt to state, in purely anatomical terms, necessary and

14. I agree with Wollheim's (1968, 1980) view that representational seeing is phenomenologically bifurcated in a way in which the seeing of aspects or physiognomy (usually) is not (contra, for example, Aldrich ((1978) p. 214) who appears to think that the same *kind* of duplicity obtains in both kinds of perception). However, I disagree with statements to the effect that we do not perceive the 'sustaining features' involved in seeing-as at all. I am inclined to think that if such features are to fulfil the explanatory function indicated by the expression 'sustaining features', then we must perceive them at least in the weak sense that, for example, we are visually or auditorily affected by the morphological or phonemic properties of facial expression and speech. No doubt we see or hear the latter only as physiognomically or semantically qualified in the stronger sense that the physiognomic and semantic features are the only ones we *notice*.

sufficient satisfaction conditions for physiognomic predicates. For the extension of any anatomical or morphological predicate suited to this purpose could not be determined without alluding to the physiognomic concept to be defined. For example, one might attempt to define a predicate, '... s-moves', true of all and only people who smile happily. '... s-moves' would then specify the form or shape of a facial movement, but all we can say with precision about the type of movement in question is that it is the type of movement produced by those who smile happily. A suggestive analogy is to be found in the relation between sound and meaning. Take the predicate, '... is the word "fool"'. This is true of a token utterance if and only if it meets certain phonological conditions. The latter, however, cannot be stated phonetically.¹⁵ The second feature of physiognomic perception can be roughly indicated by noting what either a representationalist or a direct realist might say who sympathised with the general notion of physiognomic perception. According to the former, physiognomic properties are properties which qualify visual sense-data in addition to properties of colour and shape (cf. Jackson (op cit.) p. 22). According to the latter physiognomic properties would be properties of the external surface of objects - specifically, of human faces. The contrast between ordinary and physiognomic properties, on the other hand, is an approximate isomorphism of the contrast between objectivist and phenomenalist conceptions of, say, colour. As such it is ontological. Thus, if we accept an objectivist account of the latter, we are committed to holding, approximately, that we see objects as coloured in various ways in part because of the colours they have. But the direction of explanation is reversed for physiognomic properties. Objects have the physiognomic properties they do because of the way in which we see them.

5.2 The outstanding question at this stage is : what do we see expressive behaviour as? or, equivalently, what are physiognomic

15. In this respect only I would concede that the application by anthropologists, such as Birdwhistell, of a "linguistic model" to non-verbal expressive behaviour, represents a genuine advance from Darwinism. The contrast is between Birdwhistell at, say, ((1971) pp. 80 & 99f), where he contrasts 'kinemes' to 'kinemorphs', etc., and Darwin at p. 202 of the 1965 edition of *The Expression of the Emotions in Man and Animals*, where he attempts to explain anatomically the *observable* difference between a genuine and an artificially induced smile. Of course, there is an anatomical difference, but we cannot assume that any "smile" produced by galvanisation will differ in the same ways from smiles produced spontaneously.

properties? In Chapter 1, I argued that a physiognomic property, such as anger, just is a property of being expressive of anger. So it seems that when we see a scowl as angry, we see it as expressive of anger. But applying the same reasoning to the definition of expression proposed in Chapter 4 will yield the appalling conclusion that when we see a scowl as expressive of anger we *see* it as prima-facie evidence that the person scowling is angry; indeed, this is only a diluted version of the unintelligible conclusion which is forced upon us by this line of reasoning. The problem, intuitively, is that the predicate '... constitutes prima facie evidence that p' cannot be *seen* to be satisfied except in a sense of 'sees' which is indeed far from literal. Sometimes, it is true, we say things like, 'You can see that p must be the case, given that q and r'. But such uses of the verb 'to see' are no longer tied to notions of perceptual ability; on the contrary, one could substitute verbs like 'to understand' or even 'to infer', thereby begging the question of the present chapter.

However, the principle which generated this puzzle is unsound. In general the biconditional : ' $\forall x(Fx \leftrightarrow Gx)$ ' does not entail the biconditional : ' $(A \text{ sees } x \text{ as } F) \leftrightarrow (A \text{ sees } x \text{ as } G)$ '. For example, one might recognize Napoleon on the basis of such indications as his stature, coiffure and the way his hand is held tucked into his coat. One thus sees someone as Napoleon. It does not follow, however, that one sees him as the victorious general of the Battle of Jena, although the following biconditional is true : ' $\forall x (x \text{ is Napoleon} \leftrightarrow x \text{ is the victorious general of the Battle of Jena})$ '. The same point applies even if the biconditional in question is true by definition. Thus one might see a certain shape as a circle. A circle is defined as a bound figure whose circumference is equal to its radius multiplied by π and then squared. But seeing something as a circle does not involve seeing it as having a circumference equal to πr^2 .

An explanation of the failure of inferences of this kind is suggested by Jackson in the context of a discussion of epistemic seeing or seeing-that ((op.cit.) pp. 155-159). Jackson maintains that 'A sees that x is F' is sometimes equivalent to 'A sees of x that it is F' or as I prefer, 'x is seen by A to be F', which is more idiomatic. ¹⁶

16. Jackson is, of course, applying Quine's analysis of statements about belief to seeing-that. For the analogue of the first transparent construction, see Quine (1956). For the second, Quine ((1960) p. 149).

The latter formulations are valued for the referential transparency of the x-position, but Jackson points out that they are still 'opaque modes of containment for general terms', ie. replacements of 'F' ((ibid.) p. 157). In other words, co-extensive predicates are not intersubstitutable *salva veritate* within the F-position. This need hardly surprise us when we consider that the positions represented as dots in 'A sees x as ... ' and 'A sees of x that x is ... ', can accommodate singular as well as general terms and, if the former, are certainly opaque (cf. Quine (1960) p. 151). If I see someone as Napoleon, it does not follow that I see him as the victor of the Battle of Jena. Admittedly, our problem is that ' ... is an expression of F-ness' is not merely co-extensive with a complex predicate involving the concept of prima-facie evidence, but is defined thereby. It might be suggested that the relevant locutions, though basically opaque as regards the second position, do sustain substitution of interdefinable terms. However, I see no justification for this claim. It ill accords with the intuitions of the previous paragraph and is plainly false if extended to familiar instances of opacity, such as that introduced by belief-statements.

The difficulty encountered by the claim that we see behaviour as expressive resulted in part from a failure to appreciate the role of physiognomic perception in the understanding or interpretation of physiognomic behaviour. Physiognomic perception and physiognomic interpretation are not to be identified; rather, the former is what we rely on when we try to achieve the latter. In general the way things look, or what we see them as, contributes to the explanation of why we take things as evidence supporting our beliefs about the world. But we cannot specify the way behaviour should look if we are to take it as p.f. evidence for attributing F-ness to its agent except in terms that are bound up with the concept of F-ness. The crucial point is that the evidence supplied by expressive behaviour is 'imponderable' (cf. Wittgenstein (op.cit.) p. 228). Imponderability is not the same as incommunicability, but there is a connection between them. Evidence which cannot be measured ("weighed"), also cannot adequately be described by means of the idioms we use for describing what can be measured. For example, a friendly smile will appear to be insincere if it is cut off too rapidly or if it is rigidly sustained, fixed, for too long. But we cannot specify a duration of n seconds and a duration of n' seconds such that a smile which endures for less than n or more than n' seconds will look

insincere. All we can say is that a smile that is too brief or too protracted looks insincere, and that a smile is too brief if and only if its duration is less than that required for an appearance of genuine friendliness. Physiognomic properties, ie. the properties in virtue of which we take behaviour as p.f. evidence for emotions, are perceptual but imponderable.

5.3 At the close of §2.2 I indicated the need to distinguish statements to the effect that knowledge about people's emotions is perceptual from statements to the effect that knowledge of the meaning of behaviour is perceptual. I objected in Section 3 to the 'seeing-in' construction on the ground that it failed to keep this distinction on display and consequently held out the false promise of facile, reductive solutions in the epistemology of intersubjectivity. It would seem that these pitfalls can be avoided at this stage by elucidating knowledge of psychological states on the basis of indirect perception and knowledge of physiognomic "meaning" on the basis of seeing-as :

(a) We perceive the emotions of others in virtue of perceiving their expressive behaviour.

(b) We perceive the expressive behaviour of others as expressive.

However, in thus separating two components which were confused in the idiom of 'seeing-in' we seem to have sacrificed the possibility of epistemological gain. The 'seeing-in' thesis was needed to support the claim that the emotions of others are in some sense open to observation. Of course, this will not meet the challenge of solipsism which could be based on the possibility of universal physiognomic illusion. But it might have suggested a middle way through a field currently divided among transcendent realists, reductionists and anti-realists with respect to other minds. A non-transcendent realism of the kind proposed by McDowell (1978) seemed plausible as an account of the acquisition of psychological concepts or the exhibition of competence with such concepts. I return to this theme in the final chapter. My concern here is to assess the following statement : 'If S expresses F-ness and is not F, then A, a witness of S's behaviour, is subject to the illusion that S is F'. The truth of this statement is a necessary condition of the truth of maintaining that we observe the mental states of others when they do feel the emotions they express. In order to establish the possibility of the right kind of physiognomic illusion we need to re-integrate seeing-as

and indirect perception without, however, confusing them anew.

The claim that we see behaviour as expressive will not generate the right kind of illusion by itself. True, there is scope for illusion here. I might see a facial expression as expressive of joy when in fact it is expressive of melancholy. What determines the real expressive function of the facial configuration is in part the way other people tend to see facial configurations of that kind. But the fact that I see a face as expressive of joy when its owner is sad does not amount to an illusion given the material assembled so far. For in this case the face may in fact be expressive of joy. A straightforward appeal to the indirect perception of emotions will not help either. For, irrespective of how the face appears to me, what I see in virtue of it is sadness just if sadness is what caused the face to take on that appearance.

The difficulty is two-fold. In order for an illusion to occur there must be a single object perceived (more or less directly) which must appear to have properties which it does not have. But so far we have been considering two objects, the emotion and the behaviour, one of which is perceived in virtue of the other. Furthermore, physiognomic properties are not psychological properties. Accordingly, seeing behaviour as expressive of anger when its agent is not angry cannot constitute an illusion, there being no object here which is such that it is not what it appears to be. One reason for *expecting* a solution to this problem (rather than concluding that other people's mental states just are unobservable) is the possibility of generating precisely analogous puzzles in other areas of perception. Thus if the only properties of immediately perceived objects are properties of colour, shape, etc., then the only illusions we can explain are those in which objects appear to have different colours, etc. from the colours they have. It is impossible to explain, for example, illusions of physical objects generated by holograms. For holograms are relatively insubstantial bearers of the phenomenal (including, perhaps, physiognomic) properties which they appear to have and they are the objects we see, albeit with a somewhat indeterminate degree of immediacy. Ordinarily holograms are counted as illusions generated by artifice, but we have been unable so far to locate a discrepancy between what is and what appears. In order to do so we need to admit within the phenomenal sphere properties not traditionally regarded as phenomenal (cf. Chapter 2, §4.2). For example, a holographically

generated image of a human being is a thing which looks human. This appearance of humanity is, I suggest, irreducible to a conjunction of classical phenomenal properties. Any attempt so to reduce it, eg. by means of a comparative use of 'looks' (Jackson (op cit.) pp. 31-33) will again undermine the illusion. The object in question does not merely look *like* a human being; it looks human. ¹⁷

5.4 The solution to the problem introduced in the previous paragraph, the problem, that is, of how there can be a discrepancy between behaviour and emotion which is aptly described in terms of illusion, involves two fairly easy stages. First we need to reconsider the question of the object perceived. There is an obvious candidate for the role of perceptual object common to the perception of emotion and of behaviour. What we perceive are the human beings, the relation between whose behaviour and emotions is in question. Then the first step towards a theory of the perception of behaviour is to state that we primitively perceive human beings as human beings (cf. Cooke (1969)). We do not perceive them as "senseless bodies", nor do we perceive them as indeterminate with respect to the distinctions between the animate and the inanimate. The claim made by Cooke and others is that even as infants we perceive adults as human. Indeed Max Scheler suggested that infantile perception of the world retains the traces of a primitive animism (Scheler (1913) Part III, Chapter II); one might appeal to some such notion in explaining the continuity observed by Wollheim between the perception of expressive behaviour and the emotionally coloured experiences we sometimes have of inanimate things (cf. Wollheim (1966) p. 95). Of course, in attributing these perceptual capacities to the infant we do not attribute to him the capacity to distinguish *conceptually* between the animate and the inanimate, the human and the non-human. Rather such attributions constitute an attempt to describe the perceptual basis (in a phenomenological rather than an empirical sense) of the *acquisition* of those conceptual abilities.

Human beings are not, of course, things we perceive directly any more than are purely physical objects. We perceive them in virtue of perceiving the surfaces of their bodies. But if we perceive human

17. To be really persuasive the point would need to be expanded considerably beyond what is compatible with the balance of this chapter. Still, in line with Chapter 2, it could be rephrased by saying that the comparative use of 'looks' is not a way out of the confines of phenomenal space; it describes a form of interpretative perception.

beings indirectly (in Jackson's sense), nevertheless, we perceive them as human beings. And if we perceive the surfaces of their bodies more directly, we perceive them as parts of human beings and not as parts of automata. The elements of physiognomic perception and the observation of psychological states can now be based in the perception of human beings in such a way that their dependence upon each other can be disclosed without reducing one to the other. When behaviour is perceived as expressive of F-ness then a human being is seen to be expressing F-ness. When the agent is also F, then he is seen to be F. But, precisely because of the additional condition just mentioned, the following bi-conditional is false :

- (a) S is seen by A to be expressing F-ness \leftrightarrow S is seen by A to be F.

If A is not F, then he cannot be seen to be F but he can be seen to be expressing F-ness. This is so because the form '... sees x to be F' carries a success grammar. It differs from seeing-that only in containing x referentially.¹⁸ However, the following conditionals are true :

- (b) S is seen by A to be expressing F-ness \rightarrow S seems to A to be expressing F-ness.
 (c) S is seen by A to be F \rightarrow S seems to A to be F.

Finally, the following bi-conditional is true (but for minor reservations to be ignored here) :

- (d) S seems to A to be expressing F-ness \leftrightarrow S seems to A to be F.

The main possible source of resistance to (d) would be a confusion of how things seem and how we believe they are. But the truth of (d) does not depend on truths about A's beliefs. It concerns those of his visual experiences which may or may not give rise to beliefs about their objects.

(d) is all that is required to justify the claim that to express emotions which are not felt is to produce the illusion that they are felt (without necessarily encouraging false beliefs). But the falsity of (a) is sufficient guarantee against any reductive consequences that might be thought to follow. Reductionism, then, is

18. I should add that I do not take the 'sees-that' construction or its referentially transparent counterparts to entail belief-ascriptions. In this I follow Hintikka, but not Howell (references given on p.96n) and not Jackson, although there is a curious concession on p. 159 of the work cited.

not justified by the allegation that we can sometimes observe the emotions of others by (and only by) observing their behaviour. On the other hand, solipsism is not by any means refuted by such a claim.

VI. CORRELATION AND APPROPRIATENESS

1. Performance and Correlation

1.1 When we say of a type of behaviour that it is an expression of a type of mental event we are committed to saying that within the community or species in question every token of that type of behaviour is an expression of that type of mental event. But if we say that a behaviour-type, ϕ , is an expression of an affect-type, F-ness, then are we committed to the statement that for every token of ϕ there is a token of F-ness? It seems obvious that we are not. If ϕ corresponds to the description : 'expression of F-ness', the correlation is imperfect, given the non-relational concept of expression defended in Chapter 1, since people sometimes express emotions they do not feel. If ϕ corresponds to a physiognomic or a morphological description, the correlation fails because people feign anger by scowling and pleasure by smiling. Might we nevertheless be committed, when we say that ϕ is an expression of F-ness, to saying that for every token of F-ness there is a token of ϕ ? Again it is obvious that we are not, but this time it will be for different reasons depending on the level of description to which ϕ corresponds. The universal correlation in the direction now under consideration fails if ϕ corresponds to the description : 'expression of F-ness', because people sometimes inhibit the expression of emotions they feel. It fails if ϕ corresponds to a physiognomic or morphological description for the same reason and because there may be different means available for the expression of a given type of emotion. Anger, for example, is expressed by fist-shaking, offensive manual gestures, verbal abuse and inarticulate shouting. Both the diversity of behaviour patterns expressive of anger and the physical impossibility of certain combinations of them - such as the first and the third with the second and the fourth of the examples given - rule out the possibility of anger invariably giving rise to instances of any one morphological or physiognomic kind of behaviour. The obvious conclusion justified by these observations is that expression cannot literally depend upon 'constant conjunction' - a phrase which usually connotes a perfect symmetric correlation between items of different kinds. For while some types of behaviour are expressive there are no perfect (even asymmetric) correlations between tokens of them and tokens of what they express.

1.2 Nevertheless we take it for granted that there are positive albeit less than perfect correlations between expressive behaviour and what it expresses. This raises a hypothetical question as to whether a type of behaviour could be an expression of a type of emotion if there were no correlation or even if there were a negative correlation between their tokens. In other words, could ϕ -ing be an expression of F-ness if for every token of ϕ there is no token of F-ness or for every token of F-ness there is no token of ϕ or both? A negative answer to this question would establish that some degree of positive correlation is necessary for expression. An affirmative answer would settle the question left open in Chapter 2 : correct interpretation of expressive behaviour is not determined by its standardly true explanation.

Wollheim has suggested the following hypothesis :

We could imagine laughter becoming universally the behaviour of those who were sad, without our thinking that laughter expressed their sadness : for there is surely no absurdity in supposing there to be pathognomic disturbance on a massive scale. (Wollheim (1964) p. 284)

This does not in fact entail either of the correlation failures under consideration, Wollheim's primary target here being the view that constant conjunction is a sufficient condition for expression. It will be clear from the preceding chapters that I share Wollheim's view on this point. It is possible, in principle, that the members of a group might all behave in a certain way when and only when they are in a certain affective state and yet consistently fail to interpret each others' behaviour accordingly. Furthermore, on the natural reading of 'laughter is universally the behaviour of those who are sad', it is compatible with the claim that laughter is also universally the behaviour of those who are amused or delighted. Formally : ' $\forall x : (Fx \rightarrow \phi x) \ \& \ (Gx \rightarrow \phi x)$ ' is not only internally consistent, but is consistent with ' $\forall x : \sim (Fx \ \& \ Gx)$ ', the last proposition being designed to capture the intuition that people are not usually sad and delighted simultaneously . Nevertheless, there are persuasive objections to Wollheim's claim which, if decisive, also indicate grounds for insisting on the necessity of some degree of correlation between expressive behaviour and what it expresses.

1.3 First, it is arguable that the passage quoted misrepresents or underestimates the bearing of phylogenetic change on what we are prepared to say about expression. The phrases 'becoming universally the behaviour ... ' and 'pathognomic disturbance on a massive scale' can

be read as implying a widespread evolutionary upheaval. The claim thus seems to be that while laughter is widely associated with mirth or delight as things are, it could come to be associated with sadness without its expressive significance being altered as a result. Strictly this is probably true but it contains an implicit and slightly misleading emphasis. A minimally less speculative analogy is available. Konrad Lorenz has maintained that the phylogenetic origin of laughter is to be located in behaviour indirectly associated with aggression or hostility; behaviour, that is, the function of which is appeasement or neutralisation of hostile tension (Lorenz (1963) pp. 152f in translation). It is, of course, an unjustifiable simplification of Lorenz's view to allege that laughter evolved from an expression of hostility but the crucial point here is that there is an evolutionary story relating laughter as an expression of amusement to some earlier, pre-humorous phase. The same story, whatever its details, presumably also explains how laughter came to be correlated with amusement, for it is difficult to make sense of the claim that laughter might have *become* the expression of amusement without the development of a correlation between them. True, the latter development could occur without the former and so Wollheim's conjecture is basically sound. But the establishment of a correlation has a positive explanatory relevance to expression.

This is further confirmed by the consequences of expanding the hypothesis in question. Suppose that the result of the pathognomic disturbance is not only that people laugh whenever they are sad but that they laugh when and only when they are sad. In other words, they cease to laugh when amused or intensely pleased. In this situation laughter would no longer be an expression of amusement even if it continued to be (incorrectly) interpreted as such. For while people might continue to take laughter as prima-facie evidence of amusement it would no longer *be* prima-facie evidence of amusement and would no longer be behaviour the exhibition of which is partly explained by the fact that it is taken to be prima-facie evidence of amusement.

However, the notion of a pathognomic disturbance is more naturally understood as involving a deviation from a norm rather than mere change. But the concept of universal or almost universal deviation from a norm is surely incoherent. In its defence an appeal might be made to the inadequacy of purely statistical criteria of normality. Normality cannot be determined by the calculation of averages; if it

were, it would be normal for human beings to have slightly less than two legs. Nor is normality determined by what is true of the majority of members of the class in question; if it were, then the normal human life-span would be diminished during periods of excessive violence. The second point, however, contains an obvious fallacy bound up with its failure to specify the relevant class. Normality of life-span may be determined by the majority of human lives. This will not be modified by the abbreviation of most lives in one or two generations.

Analogously the notion of a massive pathognomic disturbance is to be construed in relation not to a single class of individuals, but to two classes one of which is a sub-class of the other. The norm is defined relative to the inclusive class which, in this case, is to be identified as the human race. The deviation may be universal within a sub-class of the human race : a particular generation or society.

There is, therefore, no incoherence, but there remains a serious risk of vacuousness. If the deviation occurs throughout a given generation and is such that members of that generation do not laugh when they experience the emotions which currently give rise to laughter but do laugh when they experience entirely different emotions, then, while laughter would not have become the expression of the latter, it would surely be pointless to insist that it remained the expression of the former. An appropriate phylogeny is not sufficient, for the reasons given earlier, to justify a physiognomic interpretation in the absence of a sustained correlation. The distinction between change and deviation, difference and deviance is drained of content.

The point cannot be met by adding that members of the deviant generation continue to interpret or to see laughter as expressive of joy. Where laughter is no longer correlated with joy but continues to be interpreted as expressive of it by members of the generation in question intuition supports the conclusion that laughter is universally misinterpreted and not that it is correctly interpreted and, as such, tends to mislead. Similarly, if all the members of a given society laugh when and only when sad and never when happy or amused, and if the relevance of correlation to expression is denied altogether, then it is not clear on the basis of the material assembled so far what would determine whether they express sadness by laughing or express happiness by laughing and laugh in order to simulate joy and disguise sadness.

2. Performance and appropriateness

2.1 Phylogenetic norms are, therefore, irrelevant to the determination of expression in generations or communities which deviate from them wholesale. But if correlations or norms within those communities are equally irrelevant then it must seem that expressiveness is perfectly autonomous with respect to behavioural regularities of any kind. At this point an appeal might be made to a concept of appropriateness. The relevant concept is not, of course, captured by saying that ϕ -ing is an appropriate thing for x to do if x is F . For this would be true of any rational action in relation to its motive. The point is rather that ϕ -ing is an appropriate thing for x to do if x is to offer prima-facie evidence that he is F , or, more briefly, is to communicate that he is F .

If appropriateness is to furnish a concept of expression which is genuinely independent of behavioural regularity there are two available models : the linguistic and the iconic. The linguistic model will allow us to say that ϕ -ing is an expression of F -ness although no-one ever ϕ -s. ϕ would thus be compared to the unuttered sentences of a language, one of which, σ , may mean that p and thus be appropriate for expressing the belief that p . The intuition that unuttered sentences have determinate meanings is part of what underlies the investigation of semantic structure. The formal theory of meaning for a language assigns meanings to all its infinite sentences on the basis of the fewest possible rules of derivation and a finite set of axioms. It may seem, therefore, that not only is it possible to specify meanings for unuttered sentences (or unused sentences), but also that sentences which are commonly used metaphorically (eg. 'You have hit the nail on the head') retain their structurally determined literal meanings. But none of this would make sense without the notion of semantic structure. In the hypothetical communities in question laughter could not be a "literal" expression of delight and a "metaphorical" or, more intelligibly, an ironic expression of sadness unless laughter could be decomposed into repeatable components with specifiable 'semantic' functions. Some social anthropologists do indeed believe that non-linguistic communicative behaviour is structured in such a way as to be comparable to language (eg. Leach (1972) p. 317f, Birdwhistle (1971)). But these analogies depend for their plausibility on a blurring of distinctions between different levels of structural investigation pertinent to language - phonological, lexical, semantic and syntactical. Furthermore,

although it is reasonable to insist that sentences retain their structurally determined meanings even when they are used only metaphorically, nevertheless they also acquire a "metaphorical meaning" which determines what they are expressions of. 'You have hit the nail on the head' is not an expression of a belief about the impact of your behaviour on the top of a nail, but of admiration at a particularly apt or acute observation.

2.2 Iconicity is not much better off in this connection. First, if there is no token of the type, ϕ , then ϕ is not an iconic expression of anything. For there is no limit on the subjects ϕ could iconically be "about" in advance of the performance of ϕ -tokens in actual contexts. This is so even if iconicity is straightforwardly definable on the basis of resemblance; it is exaggerated if iconicity is construed in terms of some looser concept applicable to representation or ideograms. Still, the hand-to-mouth gesture discussed in Chapter 4 might be an expression of the desire to eat without ever being associated with it. The gesture might always and only be used in feigning hunger. So while iconicity could not assign expressive properties to unperformed types of behaviour, it could explain the expressivity of behaviour sometimes exhibited but never in conjunction with what it expresses. Iconicity does offer a genuine escape from correlation but it is inapplicable to facial expression for a more fundamental reason : facial expressions simply do not resemble the emotions they express. Laughter has no visual or auditory properties in common with delight or amusement since the latter lack visual and auditory properties altogether. Nor does laughter resemble some other behaviour pattern which delight involves a disposition to exhibit; if delight involves any disposition it just involves the disposition to laugh or smile. Stuart Hampshire's theory is therefore inapplicable at this point (cf. Chapter 3, Section 3 and Chapter 4, Section 2). Finally, some iconic communication succeeds by producing a configuration which resembles in a salient respect the object or state of affairs to which it refers (cf. Bennett (1976) p. 139). For example information about the approximate length of an object can be transmitted by means of a gesture producing a corresponding distance between the hands. But it is far from clear what could be meant by saying that laughter matches the type of situation which standardly elicits it. Indeed, the relevant type of situation can only be specified as amusing or delightful, thus begging the question.

2.3 The claim under consideration, namely, that ϕ -ing is an expression of F-ness just if ϕ -ing is appropriate for the expression of F-ness, could be protected from the criticisms of §1.4 only by manoeuvres which either trivialise it, or render it dependent on exactly what it has been introduced to replace. Thus an appeal might be made to the reflexivity of resemblance but this would result in a vacuous parody of Hampshire's account of imitative iconicity and would, in addition, depend on the holding of a correlation between behaviour and what it expresses. Indeed the holding of such a correlation between behaviour and emotion might be thought to render the former appropriate for the expression of the latter. This is obviously irrelevant to the attempt to replace correlation as the (partial) basis of expression with appropriateness though it may bear on the more plausible claim that expression is characterised by a kind of appropriateness in addition to correlation. However, the holding of a correlation does not generate expressive appropriateness unless it is observed to hold by those who interpret instances of the behavioural correlate and this explains their ability to interpret correctly.¹ This suggests a more plausible and definitely non-vacuous claim: what makes ϕ -ing appropriate for the expression of F-ness is the fact that ϕ -ing is nearly always interpreted as an expression of F-ness. This, suitably relativised to communities, generations or species, is true and constitutes a genuinely autonomous concept of appropriateness as opposed to correlation. But for the reasons given in §1.3, it is not true that ϕ -ing is an expression of F-ness just if it is in this sense appropriate for the expression of F-ness. The way a type of behaviour is standardly interpreted by those who observe tokens of it is insufficient to determine what it actually expresses even if it does make the behaviour in question appropriate for the expression of what it is interpreted as expressing. For the same reason it may be that a kind of subjective iconicity or correspondence (cf Wollheim (1966)) characterises the habitual and unreflective interpretation of other people's physiognomic behaviour, in which case the relevant concept of physiognomic appropriateness is needed to supplement but cannot replace the concept of correlation in the analysis of expression.

1. I return to this theme in Section 4.

3. Correlation as a necessary condition

3.1 Even if one accepts that for certain kinds of expression - linguistic, iconic - the appropriateness of ϕ for the expression of F-ness is a sufficient ground for asserting that ϕ is an expression of F-ness even in the absence of any correlation between them, this will not help with facial expression. The linguistic and iconic models of appropriateness are inapplicable, with the consequence that the dilemma of §1.4 remains unsolved. The dilemma was produced by the allegation that in a given society laughter could be the expression of amusement while coming to be correlated with some entirely different and incompatible emotion. There is no adequate phylogenetic, structural or iconic explanation of the expressivity of laughter in such a community. The elimination of these alternatives compels the admission of some degree of correlation as necessary within even deviant communities if it is true of *them* that laughter is an expression of amusement, delight, etc.. In fact this conclusion is not entirely *faute de mieux* but is indicative of important differences between facial expression and linguistic or iconic methods of communication which constitute the significance of the claim that the former is a kind of natural expression. The extension of the latter concept is determined primarily by ontogenetic considerations. The explanatory relevance of phylogeny is established only when certain ontogenetic conditions have been observed to hold. A crucial component in the concept of the natural expression of an emotion is that of a disposition or tendency to behave in a certain way whenever that kind of emotion is experienced. The imputation of an expressive tendency or disposition to a class of individuals does not, of course, entail the re-admission of a constant conjunction as such but it does require some degree of correlation between behaviour and what it expresses. For while a disposition may be inhibited or suppressed on some or even most occasions of its activation, it is arguable that were such a disposition invariably inhibited, there would be no ground for its attribution. However, reasons for doubting this claim will be mentioned in § 3.3.

3.2 The concept of a tendency among members of a community to laugh when amused takes the form of a generalisation about the members of the community in question. Each member is inclined or disposed to laugh when amused even if he sometimes inhibits that inclination. Is it possible that every member of the community always inhibits the inclination to laugh whenever he is amused? Intuition seems to support

an affirmative reply, but at the same time it is difficult to see how to verify such an affirmation (except in a peculiar case to be discussed shortly). One response to this puzzle is to focus on the notion of a suppressed tendency. This notion, it might be contended, is tantamount to the concept of a tendency which was manifested even if now it is suppressed. In terms of inclinations, this amounts to the claim that members of the community laugh when and only when they are amused at least until they have learnt to inhibit (or feign) the inclination to laugh. In fact this is all that is required for the purposes of natural expression; it does not entail the generalisation that people inhibit an inclination to laugh on every occasion of feeling amused. Habitual inhibition often leads to total suppression of the original inclination; many adult males are never disposed to weep when miserable, but they have usually passed through a phase during which tears had to be suppressed by a deliberate exercise of self-control whenever sufficiently acute distress was experienced.

This line of thought suggests that ϕ -ing is a natural expression of F-ness among humans only if ϕ -ing is actually correlated with F-ness at a phase in the development of each individual in every human community. More precisely, ϕ -ing is a natural expression of F-ness only if everybody ϕ -s when and only when he is F at least until he has learnt to control the activity of ϕ -ing, to ϕ or refrain from ϕ -ing at will. The difference between natural and other forms of expression is in part the difference between behaviour which we learn how to inhibit under certain conditions which dispose us to behave in that way and behaviour which we must first learn how to exhibit and then become disposed to exhibit under suitable conditions (cf. Danto (1973) pp. 163f.). Evidence that a given type of behaviour is naturally expressive of a type of emotion, expression of which is suppressed, will consist in evidence which supports a bilateral correlation between behaviour and emotion among infants up to a certain age if not among adults.

It is worth contrasting this ontogenetic requirement to the inadequate phylogenetic proposal considered earlier. According to the latter, ϕ -ing could be an expression of F-ness in a deviant community or population, P, even though ϕ and F-ness are not correlated in P, because ϕ is correlated with F-ness in a larger class, humanity, of which P is a sub-class. According to the new ontogenetic requirement, ϕ -ing could be an expression of F-ness in P, despite the absence of a correlation in P as a whole, so long as there is a correlation between F-ness and ϕ in a sub-class of P, ie. in its infantile population. If, however, in

the allegedly deviant community, not even infants conform to the required regularity, then there is no justification for the assertion that ϕ -ing is an expression of F-ness within it. This in turn would threaten the claim that ϕ -ing is a natural expression of F-ness in the species of which P is a sub-class. If P corresponds to a generation, then we would be strongly inclined to say that laughter was but is no longer an expression of delight among human beings. If P corresponds to a society, then it constitutes prima facie evidence against the claim that laughter is a natural expression of delight among humans and for the claim that laughter is a conventional or culturally determined expression of delight among some humans who are not members of P.

These reflections encourage us to try for a modified version of the thesis that correlation is a necessary condition of natural expression. In fact such a condition might be expected to have a wider relevance, being applicable not only to natural expression, but also to conventional but unstructured methods of communication typically involving gestures. For example, the manual 'V-sign' which is the reverse of the sign for victory, is taken by most English speakers as prima facie evidence of hostility on the part of its perpetrator, but many of those who grasp its significance would nevertheless refrain from using it because of a sense of the gesture's vulgarity. Thus the condition we are looking for might best be introduced as a condition on non-iconic, non-linguistic forms of expression somewhat as follows :

ϕ -ing is a non-iconic, non-linguistic expression of F-ness in a population, P, only if there is a non-empty sub-class of P, C, such that anyone who is a member of C ϕ -s if and only if he is F & most members of C sometimes ϕ .

If ϕ is replaced by a verb of facial expression then C will probably include all and possibly include only infants in P. If ϕ is replaced by a verb which attributes a conventional - eg. manual - gesture to an agent then C will only include members of P who have had time to learn and do not object to the use of the gesture.

3.3 There are obviously several problems with the proposed condition as formulated above. First, as it stands it does not exclude very much at all. Suppose that just one member of P, Fred, sneezes when and only when he is afraid. C could be designated as the class of members of P who sneeze when and only when afraid and would thus contain just one

member, Fred.² Admittedly the condition introduced in the preceding paragraph is meant only to be necessary, but its capacity to accommodate this example suggests that it is rather trivial (though not completely trivial since it still discounts cases where sneezing and fear coincide on only one occasion although each occurs independently with a reasonable degree of frequency). Furthermore, if the scenario is filled out by adding that most members of P take sneezing as p.f. evidence of fear then, although Fred is the only member of P who instantiates a correlation between sneezing and fear, sneezing is an expression of fear in P. While it is not obvious that the concept of expression should exclude such a situation, it is so intrinsically odd, and so unlike any situation we are actually likely to confront, that on balance I see no harm in making the concept exclude it.

A first step would be to stipulate that C should not merely be non-empty, but should contain a substantial proportion of P. The vagueness thus introduced is no objection to this proposal. This would seem to allow for those cases with which we started out, but it is worth building into the condition a further intuitively sanctioned requirement that members of P who do not ϕ when F have some motive for refraining from doing so. The point of this requirement is that where ϕ -ing is a natural expression of F-ness, members of P who are not members of C have had to inhibit the inclination to ϕ . Admittedly, where ϕ -ing is a conventional manual gesture it would be wrong to insist that all members of P are inclined to ϕ when F, or have at some time been so inclined. But the additional requirement holds so long as the following, weaker, counterfactual claim is true : members of P who are not members of C would ϕ when F if they had no reason for not doing so. Of course their dislike of the gesture may be so strong that they do not even feel any inclination to ϕ .

The proposed condition is also open to the objection that it is too strong. The excessive strength of the condition is due to the universal quantification over times implicit in it which requires ϕ -ing only when they are F and on *all* occasions on which they are F. This seems implausible and the quantification ought to be made explicit but toned down in such a way as to specify the behaviour of members of C most of the time. Accordingly, the condition should at least be rewritten as follows :

2. This, and another counter-example introduced on p. 133 were suggested to me by Mark Sainsbury in the course of discussing an earlier draught of this section.

ϕ -ing is a non-iconic, non-linguistic expression of F-ness in P only if there is a substantial sub-class of P, C, such that for any person, x, who is a member of C, and for almost any time, t, x ϕ -s at t iff x is F at t, members of C sometimes ϕ , and all other members of P have a reason for refraining from ϕ -ing.

However, while this loosening represents an improvement, this formulation of the correlation requirement still fails to accommodate the phenomena of what may be termed 'physiognomic synonymy' (cf. §1.1) and 'physiognomic ambiguity'. Physiognomic synonymy arises where there are two alternative ways of expressing the same kind of emotion. Anger, for example, may be expressed, even by members of C, in a variety of ways none of which, consequently, will be associated with anger even on most occasions of its provocation. Conversely the possibility of physiognomic ambiguity needs to be recognized. To some extent it has been recognized implicitly in the foregoing discussion of laughter as expressive of delight *or* mirth. Delight and mirth may either be regarded as entirely dissimilar affects, or they may be regarded as variants of the same kind of affect neutrally specifiable as, say, 'euphoria'. I doubt whether consensus can be achieved over issues of this kind given the inevitable vagueness of the intuitive affective classifications available. To mention one example of a thriving controversy of this kind; among social scientists there are those who believe that smiling is equivalent to partially suppressed laughter and carries a diminished significance of the same sort (eg. Lorenz (1963) p. 152) and those who believe that smiling is on some occasions mild laughter, suggesting partially inhibited mirth or open mild amusement, and on other occasions an altogether different kind of expression, with a distinct phylogenetic and ontogenetic background and indicative of an altogether different range of affective states including friendliness, contentment and so on (eg. van Hooff (1972) pp. 212 f). It is difficult to determine what would settle the debate in a genuinely definitive way; still, some support for the view that smiling is inherently equivocal might be derived from the confusion and misapprehension to which it is known to give rise cross-culturally and even within cultural boundaries (where these can be delineated with any clarity). Either way, the possibility of physiognomic ambiguity ought to be admitted, at least in principle, if only to avoid prejudging the debate among ethologists.

Unfortunately, it turns out to be surprisingly difficult to provide a formulation of the sought for necessary condition which is not either completely vacuous or fails to accommodate physiognomic ambiguity

and synonymy. In itself I do not believe this constitutes a decisive objection to the view that some degree of correlation is a necessary condition for the kinds of expression recently discussed, but the difficulty does not seem to be sufficiently important to justify the complexity that would have to be countenanced in dealing with it at all formally. The task will seem even more futile in the light of the following, seemingly devastating objection to the claim that correlation in either direction is, in the strict sense, a logically necessary condition of natural expression.

Consider the following scenario. A species which is not unlike humanity, is such that its members begin to experience a certain emotion, F, only at puberty. When they experience F, they are (naturally) disposed to express it by ϕ -ing. Within this species there is a community, P, whose members never ϕ because, before they reach puberty, they are taught by other members of P that ϕ -ing is totally unacceptable. But the following counterfactual is true : if any member of P were seen to ϕ by another member of P, the latter would take the former's ϕ -ing as p.f. evidence that he is F (cf. n.2 above). Should we not say that ϕ -ing is an expression of F-ness in P and a natural expression of F-ness in the species of which P is a part? It seems obvious that we should and that consequently ϕ -ing could be a non-iconic, non-linguistic expression of F-ness in P without there being any correlation between them.

There are three points to be made about this example. First, it is difficult to see how to construct a similar counter-example to bear on the claim that correlation is necessary for conventional unstructured signs. I suspect that it is impossible to do so, suggesting that correlation is necessary for conventional but not for natural expression (a suggestion which contrasts interestingly with the view put forward, for example, by Alston. See section 4 below.). Where ϕ -ing is a natural expression of F-ness it seems that all that is strictly necessary is that members of the species of which this is true should be disposed to ϕ when and only when F (subject to possible refinements dealing with physiognomic ambiguity and synonymy).

Secondly, the hypothetical example described above presupposes that the deviant population possesses a language. Without a language it would be impossible (a) for the situation to arise and (b) to verify that it had arisen. (a) is true because without a language members of P could presumably only learn to refrain from ϕ -ing by first ϕ -ing and experiencing some disagreeable result on several occasions. (b) is true because the only evidence that members of P were disposed to ϕ

when F would consist in their avowals to that effect. This, in fact, presents problems for the outsider, for whom it would be difficult to discover the meaning of native expressions meaning 'to ϕ '. Such problems are awkward, but probably not insuperable.

Finally, a situation of this kind is most unlikely to be realised and most cases of non-iconic, non-linguistic expression actually encountered will satisfy the correlation requirement as stated on p. 132 (but for ambiguity and synonymy). Thus, while the condition is not strictly necessary, it does seem to capture what is distinctive of genuine examples of the types of expressive behaviour in which we are interested.

3.4 How distinctive an aspect of expression is the degree of correlation? It is, I suggest, distinctive in two respects. First, it is sufficient to contrast expression to most other, non-expressive types of action. If we retain ϕ as a variable ranging over types of bodily movement_T, F-ness as ranging over types of mental events or states including beliefs and desires as well as emotions, and introduce a new variable, ψ , to range over types of action determined by the types of events to which they give rise beyond the surface of an agent's body, then there will be very little correlation between tokens of

.../...

F-ness and tokens of ϕ or ψ for most F, ϕ and ψ . Certainly there will be no two-way correlation between F and ψ even if for many ψ (but certainly not most ψ because of branching causal sequences of indefinite protraction emanating from any ψ -token), tokens of ψ will usually be accompanied by tokens of F-ness, eg. by wanting to ψ . Correlation in the opposite direction is precluded by Davidsonian 'holism of the mental realm' even for wantings to ψ and ψ -ings.³ Any attempt to establish such a correlation would be vitiated not merely because of the possibility of suppression for its own sake, but because of the endless possible variation in the beliefs and competing desires of a human agent.

Even if Davidsonian holism is thought to be exaggerated, it is sufficient for the contrast drawn here that there is no correlation in either direction between F-tokens and ϕ -tokens for most F and ϕ . This does not depend on an extreme holistic picture but on the obvious fact that most bodily movement types will have different consequences on each occasion of their instantiation because of the variation from token to token of the standing causal conditions under which they are performed. Such variations will be registered more or less accurately as beliefs in most agents, thus contributing to the determination of different bodily movements in different circumstances where the same type of end is sought. Similar vicissitudes do not affect the relation between bodily movements and those of their results which justify calling them expressive to nearly the same extent. They do not, therefore, threaten the correlation between bodily movements such as smiles and affects such as contentment. We do not have to arrange our faces differently every time we express the same kind of emotion physiognomically.⁴

Secondly, the imposed condition applies only to natural expression and unstructured conventional signalling systems insofar as the latter may be conceived of as techniques for the expression of a limited range of beliefs. It does not apply to iconic modes of expression or to an infinite portion of any natural language comprising all its unuttered sentences or belief-, desire-, etc.-expressing utterance types. The condition, which is satisfied by facial expression, entails that ϕ -ing is not a natural (or at least, non-iconic, non-linguistic)

3. cf. Davidson ((1970) p. 217 in 1980) and also (1980) pp. 263-4.

4. The fact that the physical descriptions of facial expressions are irreducibly vague does nothing to diminish the clarity of the contrast indicated.

expression of F-ness in P if no member of P ever ϕ -s when he is F and each member of P ϕ -s only when he is not F. The total absence of correlation conjectured at the opening of §1.2 is incompatible with the kind of expression under discussion. But it is worth noting that the possibility in P of universal physiognomic dissimulation or universal physiognomic simulation is not denied, whether it is pathological and involuntary or hypocritical and voluntary. What is denied is that the physiognomic means of simulation and dissimulation would, under these circumstances, be an expression of what it simulates.

There is at this point both a contrast and a similarity between physiognomic expression and language which will help to clarify the last point. First the similarity. If there is a sentence in the language of P, σ , which members of P utter when and only when it is dark, then it would be ludicrous to distinguish two languages, L and L', which are identical in every respect except that in L, σ means that it is dark and in L', σ means that the sun is shining and then to ask whether L' is the actual language of P and their 'actual language relation' is deviant, or L is the actual language of P, the actual language relation being perfectly standard.⁵ This matches my earlier point about laughter and delight at §1.3.

The contrast arises because it is virtually inconceivable, for reasons to be spelled out in Chapter 7, that members of P could persistently utter σ when it is not dark and continue to understand σ as meaning that it is dark. But an analogous situation for laughter is conceivable, or at any rate inconceivable for different reasons. Members of P might always laugh when sad and always be misled by each other's laughter or, if not actually misled, prone to misinterpret or misperceive each other's laughter as expressive of delight. In this case laughter would neither be an expression of delight nor an expression of sadness. The concept of expression as it applies to the human face is defined by the convergence of two conceptually independent notions; correlation and interpretation.⁶ The relative independence

5. The point is a familiar one. My way of putting it is influenced by Lewis ((1968) pp. 176ff) and Peacocke (1975).

6. It is interesting to compare Wollheim's view of artistic expression as lying 'at the intersection of two constituent notions of expression' (Wollheim (1968) p. 46). One of those notions is that of natural expression of which I maintain that it too is complex and that (in contrast to Wollheim), the criteria which determine its application could, in principle, diverge.

of the constituent notions is a mark of the naturalness of facial expression; I will justify this claim in more detail in Chapter 7.

4. Interpretation and correlation

4.1 The establishment of a correlation between a kind of behaviour and a kind of mental or affective condition does not provide a sufficient ground for claiming that the former is an expression of the latter. First, though there is no correlation between most types of behaviour and most types of mental event - eg. types of desire, nevertheless, some action-types are highly correlated with associated desire-types. Indeed in some cases the correlation will be better than it need be for expression. For example, eating is generally accompanied by a desire to eat (if not by hunger or a need for nourishment), and gastronomic desires, at least in privileged communities, generally give rise to gastronomic activities. Secondly, the claim that correlation is sufficient for expression is plainly contrary to the basic tenor of the preceding chapters. However extensive, however exceptionless the correlation, the correlated behaviour will not be an expression of its emotional correlate unless tokens of it are standardly interpreted as such by members of the class in which the correlation holds.

However, theories which aim to elucidate expression on the basis of constant conjunction are unfairly characterised as attempting to define expression in this way irrespective of interpretation. Rather the aim is to represent the constant conjunctions which hold between behaviour and emotion as explanatory of the latter's interpretability and even of its physiognomic perception.⁷ A proponent of just such a theory is Alston who, in (1965), maintains that while linguistic and facial expression have in common the fact that they are taken by suitable witnesses to be a 'reliable indication' of what they express, facial expression is so construed because of a *de facto* correlation uniting it to what it expresses, but linguistic expression is so construed because of 'a general practice of *using* the sentence in a certain way' (Alston (1965) p. 21). Alston uses the term 'practice' to mean something like 'convention' or 'rule', but this leaves the contrast he has in mind somewhat obscure. In a footnote on the same page, however, he suggests that the contrast is a matter of the

7. It is the second view which constitutes the main target for Wollheim's criticisms in (1966).

direction of dependence between expression and reliable indication in each case : facial expression is expressive because it is a reliable indication; an utterance of, for example, 'I'm disgusted', is a reliable indication of disgust because it is expressive. In other words Alston's point can be made without presuppositions about the dependence of linguistic meaning upon convention thus avoiding, or at least postponing a problem in his exposition; how is the concept of a convention to be understood independently of the concept of a behavioural regularity?

The definition of expression arrived at in Chapter 4, because it is intended to be general, ie. indeterminate with respect to distinctions between linguistic and non-linguistic, conventional and natural forms of expression, left open the question of why members of P take each other's behaviour as prima-facie evidence for (or a reliable indication of) what it expresses. The difference between linguistic and facial expression now seems to be partially determined by different possible answers to that question :

- (i) Members of P take ϕ -ings by other members of P as p.f. evidence that the ϕ -er is F *because* members of P usually are F when they ϕ .
- (ii) L-speakers take utterances of σ by other L-speakers as p.f. evidence that the σ -utterer believes that p *because* σ means that p in L.

(ii) corresponds in an approximate way to Peacocke's use of the notion of prima-facie evidence in describing the relation between a population and an *interpreted* language (cf. Peacocke (1975) p. 164). Of course, much more is required to complete the definition, notably, the correct insertion of a common-knowledge condition, if not of a concept of convention (cf. (ibid.) p. 169). The confusion in Alston's article arises in part from a failure to distinguish the very dubious application of the concept of convention to the definition of the concept of an interpreted language,⁸ and the questionable but not obviously incorrect

8. Thus Peacocke, in the quoted article, accepts the widespread view that L is an interpreted language if there is a theory of truth and theory of force for L where, roughly, the former relates sentences to the world by means of truth-theorems, and the latter potentially relates them to the minds of speakers by means of a function, f, between sentences and propositional attitudes. Thus L is a fully interpreted language if for every sentence, σ , in L we have a theorem : $T(\sigma) \equiv p$ and a function : $f(\sigma) = \psi$. (Peacocke (ibid.) p. 171)

application of the concept of convention to the elucidation of the actual language relation.

4.2 However, the import of 'because' varies in relation to the nature of the explanatory project in progress. Bearing this in mind, there may be a sense of 'because' in which (i) is true and a sense in which (ii) is true, but there is no sense of 'because' in which (i) and (ii) are both true. Correlation cannot contribute to the explanation of the ability of members of P to interpret tokens of ϕ what a theory of meaning for L contributes to the explanation of L-speakers' ability to interpret token-utterances of σ . It turns out, on the contrary, that given any explanatory project we might initiate with respect to both facial and linguistic expression, either there is no contrast between them which can be described in terms of correlation and meaning or the contrast is the reverse of that suggested by Alston.

It is possible to distinguish four explanatory projects :

- (a) Radical interpretation;
- (b) Ordinary interpretation by adult L-speakers or members of P;
- (c) The phylogeny of interpretation;
- (d) The ontogeny of interpretation

A more detailed discussion of the role of correlation and meaning in each explanatory project will be postponed to Chapter 7. At this point I will confine myself to some brief remarks which suggest that Alston's distinction is ineptly applied to any one explanatory level. The distinction is introduced as an account of the different ways in which inferences from facial expression and inferences from linguistic expression are supported. As such it is plainly irrelevant to the phylogeny of interpretation (c) where the sense in which members of P interpret each other's behaviour the way they do *because* of the emotional state with which it is associated would have to be captured by a notion of *function* applicable to evolution in general.⁹

9. cf. Chapter 4 §4.3. Construed phylogenetically (i) is true, but presumably it is true also of the evolution of natural languages. Note that this is not incompatible with the statement in the definition proposed at 4 §4.3 to the effect that members of P behave the way they do because such behaviour is taken by members of P as p.f. evidence for what it expresses. The function of expressive behaviour in general is the promotion of common knowledge among members of P (or common registering), about the environment, which includes the mental states of communicators. This function is realised only if :

(a) What members of P take as p.f. evidence for psychological facts about members of P is in fact correlated with those facts.

Also, if Alston's aim is to distinguish grounds of inference then it is inapplicable to ordinary adult interpretation either of linguistic or non-linguistic behaviour (b), if the view that everyday unreflective interpretation is perceptual rather than inferential in character is taken seriously. Still, it might be insisted that interpretation involves a perceptual or unreflective analogue of inference to which Alston's distinction is applicable. This might be elucidated through the possibility of coming to see things in certain ways as a result of having become familiar with the way things are. It depends, therefore, on the application of Alston's contrast to the way in which we learn to interpret behaviour (d).

But the crucial failure of the meaning/correlation opposition is in its application to ontogeny and, for the same reasons, to radical interpretation (a). First, what distinguishes radical from ordinary interpretation is precisely the refusal whilst engaged in the former to rely upon any antecedent knowledge of the subject-matter. The radical interpreter of *language*, therefore, has nothing to fall back on if not the correlations which can be *observed* to hold between utterances of a sentence and the conditions under which they are elicited by appropriate interrogation and under which they elicit assent or dissent from native speakers. The infant learner is in exactly the same predicament vis-a-vis his first language. It is, therefore, in connection with language that the inferential procedure available to the radical interpreter and the learning processes of the infant are supported by correlation. Alston's contrast is thus undermined.

More importantly, it appears to be reversed because there is no *observable* correlation between instances of expressive behaviour and instances of what it expresses. True it might be pointed out that the same difficulty afflicts the radical linguist and infant learner with respect to the expressive part of language, specifically, the part which comprises sentences used for attributing emotions to the utterer. But here radical interpretation reflects the ontogeny of interpretative competence with self-ascriptions of psychological predicates to the extent that in each case the self-ascriptions are interpreted by correlating them with non-linguistic expressive behaviour. We learn

9. continued

(b) The behaviour which members of P exhibit when in certain states is taken by members of P as p.f. evidence for those states and not different ones.

to understand a sentence, σ , as an expression of distress by correlating it with antecedently understood facial expressions of distress. Some reflections such as this underlies Wittgenstein's concept of expressive behaviour as the *criterion* of, or reference-fixing device for the semantic determination of avowals (Wittgenstein (1955) §244). The same approach is obviously impotent to determine the meaning of the facial expressions on which we rely in determining that of linguistic expressions.

An alternative method could be based on correlations between behaviour and situations which can be expected to give rise to certain kinds of emotions. I examine the prospects for this approach in the next chapter; they are not promising for radical interpretation and are, within the bounds of psychological plausibility, even worse for ontogeny. But even without anticipating these conclusions it is clear that the understanding of expressive behaviour is not explained by the correlation between it and what it expresses. Such correlations, therefore, do not provide the material for a sufficient condition of expression. Furthermore, the argument of this section has issued in a premiss for a negative argument for the innateness of the capacity to recognise facial expression; the negative argument will be completed if it turns out that the infant is not in a position to learn what adult behaviour expresses from the conditions under which it is exhibited. I will close this chapter with an incidental positive argument for the same conclusion, relying on premisses about physiognomic appropriateness.

5. Interpretation and appropriateness

5.1 The relevant concept of appropriateness here will not be based on iconicity or structure, and its role will not be to replace correlation altogether in the analysis of expression, particularly natural expression, but to replace it in the analysis of interpretation. The concept of natural expression requires both elements : we come to interpret people's behaviour as expressive because it seems to 'match' or 'correspond' to the moods or emotions it expresses; the presence or absence of a correlation between behaviour perceived in this way and the emotions we consequently interpret it as expressing is what determines the correctness of our interpretations. I have alluded in passing and more than once to Wollheim's elucidation of physiognomic perception on the basis of 'correspondences' (Wollheim (1966)). The

kind of appropriateness or matching to which Wollheim refers is not iconicity, but it is not entirely dissimilar from it. It might be described as 'subjective iconicity'. Unfortunately, it is difficult to differentiate in a perspicuous fashion between subjective and objective iconicity; surely iconicity in general depends upon subjective standards of similarity and ultimately, no doubt, upon an at least partly innate capacity to classify experience according to certain natural kinds. Thus Wollheim writes :

The difference between iconic and noniconic signs, which is generally treated as though it were a difference in the relation in which the sign stands to the referent, is really a difference in the relations in which we stand to the sign.
(1968) p. 157)

In certain cases iconicity accrues to the relation between sign and audience simply as a consequence of the latter's familiarity with instances of the former (*ibid.*). But if this is so then we would have to describe the perceptual element in linguistic understanding as iconic (a problem noted by Wollheim at (*loc cit.*)). Still, this difficulty need not disturb us so long as we take it that the operative concept is subjective and can be distinguished from the standard examples of iconicity which, though also subjective, are partly explained by features objectively possessed by the material sign.

The question of whether or not a clear distinction can be drawn between subjective and "objective" iconicity need not preoccupy us at length. For there is a second difficulty attached to the application of the former to the perceptual basis either of linguistic or of physiognomic recognition. There is scope here for a *phenomenological controversy*. I am not persuaded that sentences in our own language or familiar facial expressions do seem to match what they express in the way in which a rural scene or even a chromatic configuration might seem to match a mood. On the contrary, such correspondences seem to depend on a degree of unfamiliarity; to many English-speakers, for example, expressions of passion uttered in French simply *sound* more amorous than their English translations. I have no idea how to settle this controversy.

5.2 These scruples do not require the withdrawal of the concept of appropriateness altogether, for the differences between the perception of correspondences and the perception of physiognomy are so phenomenologically subtle that, for all I know, there may be nothing corresponding to them in many people's experience. Whether or not this is so,

the similarities far outweigh the contrasts. Physiognomic perception and the experience of correspondences are both forms of aspect-seeing or the kind of extended seeing-as which intrigued Wittgenstein ((1953) Part II, Section XI). What we see things as sometimes depends on familiarity with what things which look that way actually are. It can be produced by an antecedently formed belief as well as contributing to the formation of new beliefs (cf. Wollheim (1980) pp. 220-1). What initially appears to me as no more than a dull, red patch with a squarish shape will take on the appearance of a decrepit double-decker bus partly as I remember that such a vehicle has always been in that place. On other occasions, it is the operation of the imagination rather than memory or intellect that determines how things look.¹⁰ Finally, Wittgenstein himself was struck by the way in which the seeing of aspects, like imagination, is to some extent subject to voluntary control ((loc cit.) p. 213). We can choose to see the duck-rabbit either as a representation of a duck or as of a rabbit.

Both physiognomic perception and the perception of correspondences differ from the foregoing cases of aspect-seeing in two salient respects. First, Ishiguro has pointed out that not all seeing-as is subject to the influence of the will (Ishiguro (1969) p. 53). Physiognomic perception is plainly involuntary; we cannot help seeing a scowl as angry, a happy smile as happy (cf. Wollheim (1968) p. 61). To some extent the same is true of linguistic understanding (cf. *ibid.*), and probably of correspondences. Those who merely affect a delicately tuned sensibility are, one suspects, incapable of seeing things in pleasingly affecting ways however hard they try to. Secondly, physiognomic perception is, in addition to being involuntary, non-rational; it cannot be influenced by beliefs previously formed as a result of a process of induction, since the required generalisation cannot in the ordinary way be established. In this respect physiognomic perception is unlike unreflective linguistic competence. It is also unlike the experience of correspondences because the latter, while comparably non-rational, is so because it is irrational. There is no criterion of correctness applicable to the emotional light in which we perceive inanimate things.

10. Some philosophers are attracted by the possibility of elucidating seeing-as solely on the basis of imagination (eg. Warnock (1976) pp. 187f). But analyses of this kind cannot plausibly be extended to physiognomic perception.

5.3 The combined effect of the involuntary and non-rational nature of physiognomic perception is to suggest that we look beyond the conscious operations of the mind for its explanation (cf. Ishiguro (1969) p. 53). When behaviour or cognitive states appear to be involuntary or non-rational we tend to locate their causes outside the areas in which reasons are formed, or at least to hypothesise that there is some causal process of a more mechanistic than rational nature engaged. When, furthermore, the phenomenon under investigation is repeated in connection with several individuals, and when all those individuals are members of an antecedently identified class such as the human race, then there is a strong suggestion that the same mechanism is operative in each case. It is a part of human nature to perceive each other's physiognomy in appropriate physiognomic ways. Since the physiognomic perception of expressive behaviour is not explained either by ratiocination or learning, or by voluntary decision, or by the straightforward operation of the machinery of perception, the case for its innate determination, for 'innately acquired receptor mechanisms' (Eible Eibesfeldt (1972) p. 306) merits serious attention.

VII. NATURAL EXPRESSION

1. Justification of the subject-matter

1.1 The aim of this chapter is to develop some of the themes introduced in Chapter 6. In the last chapter discussion remained at a fairly conceptual level in the sense that while certain suggestions were put forward concerning the nature of the evidence required for the claim that a type of behaviour is innate, and in that sense, natural to humans, the justification for these, as yet rather tentative claims, depended on questions about what we would be inclined to say under certain hypothetical conditions. In this chapter I proceed to examine the way in which the issue of innateness might be settled in the light of data widely agreed to have been gathered by anthropologists, psychologists and ethologists.

Conceptually the contrast between conventional and natural forms of expressive behaviour is a legitimate one (cf. Alston (op cit.)). One way of establishing the naturalness of much physiognomic expression, therefore, would be to demonstrate that it is not conventional. But this distinction needs to be decomposed. There are two necessary conditions of conventionality, both of which can be shown to be corollaries of Lewis's (1968) definition of convention, and neither of which are met by physiognomic expression. It is not necessary to demonstrate the connection between these conditions and convention insofar as failure to meet them constitutes a sufficient ground for innateness. The conditions to which I allude are the following :

- (a) Conventional regularities are culture-specific;
- (b) Conventional regularities are regularities to which those who conform have learnt to conform.

The first point is more or less explicit in an analysis of the kind put forward by Lewis. It is a convention in a group, P, to behave in a certain way in certain circumstances, if members of P do behave in that way in those circumstances, and they do so because other members of P do so and because they are expected to do so by other members of P. This requires that members of P should be related to other members of P in such a way as to explain the transmission of the disposition to conform to the convention. The nature of the relation which must therefore bind some members of P to some other members of P is captured

in condition (b). Members of P learn to conform to conventional regularities from other members of P. The concept of a convention, therefore, is defined in part, by what I will call 'pedagogic transmission'. Pedagogic transmission cannot occur across cultural barriers where a barrier between a group P and a group P' is defined as cultural if it is a barrier to pedagogic transmission. In other words, such contacts as there are between members of P and members of P' are not such as to generate conformity in P to conventions which prevail in P'. The notions of a cultural barrier and pedagogic transmission are cognate.

Of course, both concepts are relative to a specific regularity under discussion. There may be certain conventions to which members of P, or of a sub-class of P, and members of P', or a sub-class of P', are parties. For example, P and P' may have different languages but certain polite gestures in common. There can be international conventions. But this will not, on the whole, seriously interfere with the bearing of species-wide regularities on the naturalness of expressive behaviour; in general the hypothesis of species-wide pedagogic transmission will be contrary to plausible historical generalisations about the degree of contact between cultures with shared behaviour patterns. International conventions are a comparatively recent phenomenon.

These remarks are sufficient to indicate why cross-cultural comparisons are thought to have a role in determining whether physiognomic expression is natural or conventional. But the question is more fundamentally dependant upon ontogenetic considerations and the evidence of developmental psychology. At what age do physiognomic behaviour patterns begin to be exhibited? Ideally, as Birdwhistell ((1971) p. 32) has suggested, comparisons should be carried out between rates of physiognomic development in different cultures. This ideal is also suggested by the conceptual requirements disclosed in the last chapter, but it appears that little has been achieved in the way of systematic research in this area. Still, some progress can be made by considering the relations of theoretical priority between the currently separate areas of developmental psychology and anthropological research.

The aim of this chapter is to show that developmental psychology provides the fundamental determinants of ontogenic issues, especially the issue of whether there is any such thing as natural

expression. I will then argue that in the light of evidence which is not actually disputed - though it is frequently neglected by those who, like structural anthropologists and psychological behaviourists, would promote relativist, learning theories of *all* behaviour - the case for natural expression is very strong. In one sense this order of priorities is obvious, but the links between developmental psychology and cross-cultural research are more intricate than a simple statement of priorities suggests. Cross-cultural comparisons are insufficient to determine the point at issue in certain obvious ways. First, where members of all cultures do behave in similar ways, there are possible explanations of this which avoid the attribution of innate capacities. There is the remote possibility of cross-cultural pedagogic transmission and also the possibility, recognised by modern universalists such as Ekman but ignored by Darwin¹, of universal environmental conditions sufficient to justify the generalisation that members of all cultural groups *learn* or are *conditioned* to behave in similar ways. Secondly, and this is crucial, the question of whether there are universal patterns of *expressive* behaviour should be understood as follows: Is ϕ -ing an expression of F-ness in every cultural group? I will argue that this question is hopelessly underdetermined by the data available to anthropologists if they proceed according to empiricist assumptions and refuse to take the data delivered by developmental psychology into account.

1.2 Before initiating the argument in detail, certain objections of a general nature must be answered. It may be doubted whether it is legitimate for philosophers to participate in a debate among social scientists. In doing so one renders oneself vulnerable to charges of "armchair psychology" or "armchair anthropology". There are various points to note here. First, it is curious that contemporary philosophers tend to evince a certain reluctance to risk the first but not the second of these pitfalls. Philosophical discussion of language acquisition is often hedged about with apologies (eg. McDowell (1978) p. 128n), but not discussions of radical interpretation. The explanation of this lies, I think, in an instinctive assumption that philosophy is concerned with the order of justification while science, including the social sciences, and scientific methodology are concerned with the order of discovery (to adopt a distinction due to Santayana).

1. Darwin ((1965) p. 14f) and Ekman ((1973) p. 171).

Thus philosophers of radical interpretation do not expect the procedures actually adopted by field lexicographers to reflect the form of their philosophical reflections. The latter are intended to exhibit the justification available for the interpretation of an alien language. But it is difficult to find a counterpart to this distinction applicable to language-acquisition. There is only the order in which the infant discovers his language and this, it is felt, must be a matter for empirical psychology.

However, the importance of scruples about armchair psychologising can easily be exaggerated. Wittgenstein, Quine and Dummett, among others, cannot be criticised for their preoccupation with the acquisition of language for it is a serious requirement of philosophical theories about meaning that they should exhibit the conditions under which it is *possible* for a language to be learnt (cf. Wittgenstein (op cit.) §224). Indeed, questions about radical interpretation and questions about language-acquisition are widely supposed to have the same philosophical relevance. The verifiability of a theory of meaning and the discoverability (by the infant or the lexicographer) of the language for which it is true are taken to impose the same condition on the philosophy of language, namely that it should relate semantics to empirical observation. In fact, for reasons to emerge shortly, I believe this assimilation to be highly misleading. It is well worth distinguishing the empirical determination of meanings from the empirical determination of theories of meaning (cf. Føllesdal (1975)); the former, but not the latter, may well admit *imponderable* evidence. Finally, whether or not such a distinction is accepted, too strict a prohibition of armchair psychology is likely to be taken as a license for social scientists to indulge in equally illicit, casual philosophising in the field or laboratory. In practice, it may be impossible to fix with any precision the boundaries separating philosophy from one or other of the scientific disciplines, particularly where the social sciences are concerned. There are certain areas of overlap, including methodology and the interpretation of data. The arguments to follow touch on both of these areas which constitute the arena of the controversy which concerns us here. That controversy is not really empirical and, as intimated earlier, the stand taken here relies on empirical data only insofar as they are the subject of widespread consensus among social scientists interested in the topic. Thus Edmund Leach, a confirmed relativist, admits some universals and some

innate contribution to the determination of facial expression (Leach (1972) p. 339). Ekman, on the other hand, committed as he is to the significance of innate determining factors, nevertheless acknowledges considerable cross-cultural variation (Ekman, eg. (1973, 1980)).

1.3 Where §1.2 was concerned with the question of the philosopher's *right* to participate in the debate over the innateness of facial expression, I turn now to the question of the philosopher's *duty* to do so. To put the matter less moralistically; why should philosophers be interested in these topics? Again, apart from the obvious answer that any subject-matter offers some potential for philosophical debate, the question can be approached via reflections on the philosophy of radical interpretation. In Chapter 6 I assumed an analogy between radical interpretation and the ethnology of expressive behaviour. The latter simply is the radical interpretation of the expressive behaviour of culturally alien communities. This analogy underlies the possibility of making a detailed comparison of the procedures involved in each case. There are important contrasts. It also suggests that whatever philosophical benefits are to be obtained from radical interpretation of language are also to be found in a discussion of the radical interpretation of non-verbal behaviour.

The second expectation is further encouraged by the connection, alluded to in Chapter 1 §1.1, between the two kinds of interpretation. This can plausibly be represented as a dependence of linguistic upon pre-linguistic radical interpretation, acknowledgement of which is implicit in philosophical theories of the former. This dependence is two-fold. First, the interpretation of native utterances whose function it is to ascribe psychological states either to the speaker or to others will rely on the natural expressions of those psychological states (cf. Chapter 6, §4.2). This matches Wittgenstein's explicit preoccupation with the natural expression of pain, though it is perhaps inept to allow the expression of sensations as such; rather it is the attitudes or emotions to which those sensations give rise which are expressed. When we are in pain we express the suffering or distress physical pain usually causes.² Whatever view is taken on this point,

2. It is not entirely certain that pain is just a sensation rather than a complex syndrome which includes sensational and affective elements. Is the description of a masochist as one who *enjoys* pain contradictory? (See Hare and Gardiner (1964)). Either way the

it is a corollary of Wittgenstein's conception of natural expression as the criterion which determines the reference of psychological terms that the former is what we rely on in the radical interpretation of the latter. . . But this accounts only for a small portion of any language and it seems that the same straightforward link does not hold between non-verbal expressive behaviour and the central, assertoric part of language. There is no natural expression of belief.

However, Davidson relies on the possibility of discerning which sentences of his language the native speaker holds-true; Quine relies similarly on the interpreter's ability to recognise native signs of assent and dissent. Literally construed in terms of natural expression these assumptions appear somewhat suspect. It is by no means certain that there are universal (albeit imponderable) indications that speakers hold-true the sentences they utter. Certainly the tone of an utterance is no guide to the honesty or otherwise of utterances in a foreign language, though there may be subtle physiognomic clues discernible by a highly trained, sensitive observer. Even so, their reliability is surely doubtful. Furthermore, it is now agreed among anthropologists that the non-verbal signs of assent and dissent are culture-specific.³ The dependence of radical interpretation upon natural expression with respect to the main body of an alien language will be less direct, more diffuse. The interpreter may be expected to rely upon native displays of congratulatory enthusiasm at his successes and exasperated frustration at his failure to grasp the meaning even of the simplest sentences. Highly relevant in this connection would be the 'eyebrow-flash' which, according to

2. continued

concept of the expression of pain should be understood in terms of the expression of a kind of distress if it is granted that emotions, but not sensations, can be expressed. This seems right intuitively, but the only defence of the claim with which I am familiar strikes me as unsatisfactory (Tormey (1971) Chapter 1). The argument makes intensionality of the expressandum a necessary condition of its expressibility (see, also, Danto ((1976) Chapter 6). This appears to be ad hoc at best, and does not obviously succeed in excluding sensation. Underlying this restriction is the thesis, due to Brentano, that intensionality is the universal characteristic of the mental. But what are sensations if not mental?

3. From Darwin until the middle of this century the opposite seems to have been assumed. But the culture-specificity of assent and dissent is one of the few uncontroversial claims for relativity made in the notorious article by Weston LaBarre (1947).

Eible Eibesfeldt ((1972) p. 300), may well be a natural and universal physiognomic sign if sufficient latitude is granted in its interpretation. As Eible-Eibesfeldt puts it, the eye-brow flash expresses a 'Yes to social contact', which ranges over approval, consent, assent, greeting, etc..

Whether or not it is accepted that radical interpretation cannot get under way without some support from the natural expressive behaviour of native speakers, it is surely indisputable that certain broader philosophical interests may be served by a theory of the latter. We rely on physiognomic evidence in seeking access to the affective facts about other people. This is so particularly when we are not in a position to assess or to trust their linguistic behaviour. The ethnologist, as he initiates his investigations, seems forced to trust his own physiognomic perceptions of native behaviour. The infantile language learner is in a comparable predicament. He will be in no position to appraise the success or failure of his first, tentative attempts at speech unless he can recognise his parents' smiles as expressive of approval, their sighs of dismay as he persists in misapplying his rudimentary vocabulary. Even as adults we are prone to give credence to the physiognomy of an interlocutor in preference to the words he utters partly on the assumption that the former is less subordinate to voluntary control than the latter, an assumption which, as we become more sophisticated, we learn to qualify. It seems unlikely, furthermore, that a theory of action will render natural expression superfluous in this connection. Actions do not speak louder than words; they do not speak at all. Ordinary functional behaviour does not deliver up the secrets of its motivation, actual or feigned, to a casual glance as physiognomic behaviour does. On the contrary, it is quite probable that in starting out to explain the actions of unfamiliar agents in unfamiliar contexts we will be bound to scrutinise any associated physiognomic appearances. Certainly, human-beings are intelligible to each other only through their behaviour. My suggestion is that even if one denies that all psychological revelation ultimately depends upon natural expression, the latter is fundamental at least in the sense that it is more readily intelligible than any other kind of behaviour.

These reflections were introduced primarily to weaken the resistance which might be met by the following philosophical contribution to a debate which at first sight appears to be conspicuously

unphilosophical. But they serve also as a partial justification of the entire essay.

2. Is any expressive behaviour universal?

2.1 It is important to stress that our concern is with the question in the rubric and not ultimately with questions such as the following : are there any types of situations which are such that members of all cultural groups exhibit the same type of behaviour in them? Undoubtedly the two questions are related, but the first is not reducible to the second. Nevertheless, the second question has been thought to capture the whole empirical content of the first. There are, moreover, strong methodological grounds favouring a reduction. Since emotions cannot be observed independently of the behaviour in which they are expressed, we can observe only the expressive behaviour itself and correlate it with situations classified according to observational criteria. This is the procedure adopted by Birdwhistell ((1971) p. 31). I will argue that the classical empiricist who enters the field without presuppositions must remain impotent to form a theory of the expressive behaviour of the subjects of his investigation.

The problem is that there are three elements to be correlated and they possess a disconcerting potential for independent variation. Reductive programmes aside, we need a theory capable of correlating expressive behaviour, emotions and eliciting situations. Let us suppose that the prospective ethnologist knows of some observable type of behaviour, ie. of some bodily movement, ϕ , that he and his neighbours at home ϕ only when they are F, and they are F only in eliciting situations of some observable kind, E. We will waive for the time being the extraordinary implausibility of the assumption that the expressive type of behaviour can be described purely in terms of bodily movements and that the relevant situations can be specified without allusion to the kind of emotion to which they give rise. Let us now suppose that the ethnologist discovers that members of an alien population, P, ϕ only in an altogether different kind of situation, E'. In E-situations they do not noticeably do anything at all, or they exhibit some other type of behaviour, ψ , which is such that the ethnologist and his neighbours ψ not when they are F but only when they are F'. There are three theories available to the ethnologist to account for ϕ -ing in P. All the theories are perfectly compatible with all the observable evidence :

- θi : Members of P express F-ness by φ-ing, but tend not to feel F in E-situations; on the contrary, they are F' in E-situations and they express their F'-ness, like the ethnologist, by ψ-ing.
- θii : Members of P express F-ness by φ-ing and they are F in E-situations but they refrain from expressing F-ness especially in E-situations (and especially by φ-ing), where they often feign F'-ness by ψ-ing.
- θiii : Members of P do not express F-ness by φ-ing, and when they are F, as indeed they are in E-situations, they express their F-ness by ψ-ing.

The cultural relativist with respect to facial expression who insists, like Birdwhistell, that his theory is based on empirical evidence in a way that Darwinism (universalism) is not, commits one or both of two fallacies. Following the psychologist, Paul Ekman's usage, we can designate them as the 'display rule pitfall' and the 'elicitor pitfall' (Ekman (1980) pp. 90-91). The relativist who infers θiii from the absence of correlation between φ-ing and F-ness and the presence of correlation between ψ-ing and F-ness ignores θi and θii which, so far, have equal empirical justification (or lack of it). He thus falls into the elicitor pitfall and the display rule pitfall respectively.

2.2 These points will benefit from being spelled out in more concrete detail. First the elicitor pitfall. The fallacy involved might be summarised as follows : the cultural relativist with respect to the expressive "meaning" of behaviour is a universalist with respect to the emotions elicited by various kinds of situation (cf. Ekman (1980) p. 90). He assumes that a certain class of E-situations elicits the same kind of emotion in people the world over. Given this assumption, it follows that if members of different cultural groups behave differently in similar E-situations then (but for the conventional interference of display rules etc.), they express similar emotions in different ways. However, the initial assumption of the behavioural relativist - psychological universalism - has no more empirical justification than the assumption that these are universal expressive behaviour patterns. Indeed, the former assumption is highly implausible in certain cases; humour, for example, is probably subject to a high degree of cultural variation; we can be more or less certain that members of different classes within just one community will not always

be amused by the same material.⁴ The problem for the psychological universalist/behavioural relativist is compounded by the enormous difficulty of specifying the relevant kinds of E-situations without alluding to their capacity to elicit some emotion. This, too, is clearly exhibited in the example of humour. Whether or not humour is subject to cultural determination, the situations which stimulate it plainly do not share any "observable" properties even within the confines of our own community. Indeed, it is far from obvious that amusing situations share any properties at all other than that of being funny; witness the disastrously limited scope of the purportedly general analyses put forward by Bergson (1911) and Freud (1905).

The display rule problem is equally clear intuitively, although considerable philosophical ingenuity may be required to render it thoroughly lucid. We need something like the distinction, by now familiar, between rules or conventions which are merely regulative and those which are constitutive (cf. Rawls (1955), Schwayder (1965) p. 287, Searle (1964)). At this point I can offer only a brief and in some ways unsatisfactory account of the distinction insofar as it bears on the issues at hand. To begin with, the distinction rests upon a general distinction between classes and kinds or types of action and is, therefore, unacceptable to those who refuse to countenance unreduced intensionality in philosophical analyses. Initially, the concept of a convention is defined on the basis of a contrast between the conditions responsible for the origin and the preservation of different behavioural regularities. The notion of a regularity in behaviour is worked out in terms of a correlation between ways of behaving and repeatable objective situations originally construed by Lewis (1968) as games of pure conflict or, more importantly, of coordination.⁵ So far, then, the concept of a convention requires only the perfectly extensional notion of a correlation between members of a class of behaviour and members of a class of situation. No mention is made of the possibility that

4. In this case the larger community, eg. the British, would be identifiable by means of various criteria such as the existence of a shared language, geographical and historical continuity, etc., but not, of course, by reference to a shared propensity for laughter.

5. I am not sure whether the game-theoretical model can plausibly be applied in any detail to all conventional regularities. It is often not clear what specific *interest* is served by coordination beyond some shared, inexplicable interest in conformity.

semantically non-equivalent descriptions may be true of all and only members of the relevant behavioural class. But at least two descriptions will be true of all members of a class of expressive behaviour; each member of it will be a ϕ -ing and an expression of F-ness.⁶ Recognition of this will allow us to distinguish conventional behaviour from behaviour which is natural but subject to conventional modification. It is worth noting in this connection that Schwayder scorns the epithet 'natural' as a means of characterising non-conventional types of behaviour on the ground that language is conventional but at the same time linguistic behaviour is natural for human beings ((op cit.) p. 284). Again, this confuses levels of description. It is natural to say that p when p is the case and certain other conditions are met, but it is the linguistic conventions which prevail in a particular community which enable one to say that p by uttering a sentence in the language of the community. For similar reasons the conventionality of an expressive action type is determined by what the conventions which regulate it "stipulate". More precisely, it is determined by the concepts under which the relevant correlation is subsumed. There is a difference between the following claims :

- (i) Members of P ϕ in E-situations
- (ii) Members of P express F-ness in E-situations.

Accordingly, there is a difference between claiming that a convention is in force to which one conforms if and only if one expresses F-ness by ϕ -ing (in E-situations), and claiming that a convention is in force to which one conforms if and only if one expresses (or refrains from expressing) F-ness in E-situations. A constitutive rule or convention is one knowledge of which is necessary for knowing *how* to express a certain kind of emotion (cp. Schwayder (op cit.) p. 187). A regulative convention or display rule is one knowledge of which is necessary for knowing *when* to express a certain kind of emotion.

If the foregoing distinction, or something like it, is sound, then the following passage is, but for certain obvious equivocations,

6. Of course this need not strictly be true if there are cases like those admitted in Chapter 6 where ϕ -ing may be an expression of different kinds of emotion and F-ness may be expressible in different ways. However, it is safe to ignore physiognomic ambiguity and synonymy at this point because certain display-rules are applicable only to specific ways of expressing the kind of emotion in question and/or to some but not all of the E-situations which tend to elicit it.

a non-sequitur :

There is clinical and anecdotal evidence that, at least in Western cultures, children must learn to smile in appropriate situations. That is, they must learn how and when to smile ... ' (Birdwhistell (1971) p. 32)

In fact the first statement is almost certainly false for many situations in which it is appropriate to smile (references are supplied later); but even if it were true, it would not justify the statement that children must learn *how* to smile. The distinction between behavioural regularities the explanation of which involves mentioning conventions or rules only regulatively and not constitutively is best illustrated by example. There is ample evidence of variations in the behaviour of different communities at funerals (see, eg. Leach (1972) p. 331-2). But now, if we find that members of a group, P, tend to smile and laugh at the funerals of those to whom they have been attached, while it would be absurdly naive to infer that members of P are shockingly callous in the face of death, it would be little better to infer that their way of expressing the misery of bereavement is by laughing. Ignoring the fact that in the light of many religious beliefs the death of a loved one should be taken as an occasion for rejoicing, we would be inclined to explain the discrepancy between the behaviour with which we are familiar and that of members of P on the basis of a display rule prohibiting the overt display of sadness and enjoining its dissimulation by means of apparent jubilation. This interpretation receives added confirmation where, as in Leach's (loc cit.) example, the conventions appear to operate selectively ;

the requirement to refrain from crying is, in our home culture, more rigorously imposed on men than on women.

2.3 The thoroughly *empirical* observer apparently faces an impasse. The evidence will not permit him to choose between the theories outlined above. There is, furthermore, no reason to suppose that he could improve matters by considering the entire expressive repertoire of the people whose behaviour he is investigating. It is conceivable that a display rule might be in force which prohibited any revelation of a certain kind of emotion under any circumstances which actually elicit it. In practice, however, even fairly radical empiricists will permit the ethnological observer to enter the field with a handful of presuppositions about human nature, a certain amount of irrepressible 'animal faith' which will enable him to pick out certain

situations as elicitors of certain basic emotions. In the spirit of animal faith, then, the investigator might be permitted to assume that all human subjects experience intense fear when faced with the plain prospect of imminent violent death. Reduced to its constitutive elements the assumption includes seemingly legitimate presuppositions about the proper functioning of the subject's sensory equipment, a modicum of intelligence sufficient for the anticipation of events in the short-term future and a totally earnest and urgent desire for self-preservation. Preconceptions of this kind will enable the investigator to sidestep the elicitor pitfall, although it must be recognised that plausible preconceptions will not always be ready to hand and that so far we have no way of distinguishing genuine animal faith from theoretical prejudice, the propositions we are bound to believe because we are human from the propositions we are bound to believe because we are, say, structuralists.

Even if this manoeuvre is allowed, the investigator has still found no way of circumventing the potentially deceptive influence of display-rules which may or may not prevail. He may attempt to rectify this by arguing that certain emotions, fear of death for example, are too powerful to be inhibited for the sake of some convention. It is these emotions through which he can effect an initial, tentative foray into the affectivity of his alien subjects. However, this is not obviously true. In *The Gallic Wars* Julius Caesar expressed admiration for those of his legionaries who continued to wear an expression of stern ferocity even in death. Our investigator, however, confronted by the same phenomenon, would be obliged to suppose that what looked to him like ferocity was for the Romans an expression of fear. It seems that he, had he been a contemporary but not a compatriot of Caesar's, would have been unable to translate certain passages in *The Gallic Wars*.

Still, at this point the ethnologist may adopt a strategy currently available to uncompromising empiricists. Rather than attempt to cope with the display-rule problem or the elicitor problem singly, he will note the possibility of alternative interpretations of Caesar's respect for the appearance of his late subordinates. Does Caesar admire the corpses' physiognomy on the ground that it conceals the terror felt at, or immediately prior to the moment of death, or does he admire it on the ground that it indicates that the corpses' late owners suppressed even the feeling of fear under naturally frightening

circumstances? There is an indeterminacy here which matches an indeterminacy arguably inherent in the concept of courage which can be understood indifferently as the ability to inhibit the manifestation of fear or the ability to suppress the fear itself. Accordingly the cultural relativist might counter the accusation that he ignores θ_i and θ_{ii} by repeating the anti-realist slogan : there is no fact of the matter. There is no way of determining which of the two fallacies he has relevantly committed or which of the two theories he ignores might be true.

By the same reasoning θ_{iii} , the relativist position, is indistinguishable empirically from θ_i and θ_{ii} . This appears to undermine the very terms of the controversy between relativism and universalism, thereby supporting neither theory. But this appearance is misleading. The relativist who appeals to anti-realism with respect to the emotions felt by members of alien cultures and indeterminacy with respect to the expressive meaning of their behaviour is not committed to θ_{iii} . He is committed only to verifiable (and often verified) propositions of the form : members of P do not ϕ in E-situations. This proposition is an element in each of θ_i - θ_{iii} , but the relativist goes on to say that the further elements in each theory are undecidable in principle, and should, therefore, be relinquished. There is no further question about whether members of P feel F in E-situations and whether ϕ -ing is an expression of F-ness. F-ness just is what people feel and what is expressed in their expressive behaviour in E-situations. Empiricism leads, via anti-realism, to a 'form of behaviourism'⁷ which will justify relativism if it is found that the correlation between expressive behaviour and stimuli varies between cultures.

2.4 This line of argument is vulnerable to several objections which I will review briefly. First, it is far from certain that the anti-realist approach outlined above really succeeds in eliminating all non-empirical material. There remains the outstanding and seemingly intractable question of how to classify the token-actions and individual E-situations to be correlated. How can we determine when a native is

7. See McDowell ((1978) p. 130). McDowell imputes a 'form of behaviourism' which is distinct from 'crude reductionism' to Dummett. I cannot consider the possible difference between reductive and anti-realist behaviourism here, but perhaps it is possible to distinguish methodological from conceptual behaviourism, in which case our relativist can be represented as committed to the former like Birdwhistell. I suspect, however, that the theoretical import of such distinctions is negligible.

ϕ -ing, when, that is, he is performing what other natives would regard as a token of ϕ ? The problem arises in part because of the ineliminable vagueness of any non-physiognomic description true of every token of a physiognomic kind of behaviour (cf. Chapter 5, §5.1). The interpreter cannot be sure that any vagueness in his own classification will be matched by the vagueness of the native classifications. As regards the relevant E-situations, we have already noted the absence of any guarantee that their classification will be determined by features accessible to the interpreter. This difficulty is compounded by the probability of internal sub-classifications of E-situations. Pleasure, for example, may be induced by the spectacle of misfortune or the sight of a friend. In each case it may be physiognomically exhibited in similar ways; would it follow from this that smiling or laughter sometimes express hostility and sometimes friendliness? Surely one would be more inclined to say that smiling is generally caused by pleasure while pleasure can be stimulated by very different kinds of situations. In general, behaviourism of any kind will never be fully vindicated until it can register highly subtle psychological distinctions on the one hand, and the classifications delivered by our intuitive affective taxonomies on the other, as differences in behaviour. It is unlikely ever to meet this requirement.

Finally, the relativist who rests his case on the indeterminacy of meaning is faced with a problem which also faces Quine : under what conditions does underdetermination constitute indeterminacy? For the theory of nature is admitted by many, including Quine, to be evidentially underdetermined. Indeterminacy, however, allegedly applies only to the theory of meaning (cf. Hookway (1978) pp. 17-43). I do not propose to tackle this question, but one point is obvious. Underdetermination, though not sufficient, is a necessary condition of indeterminacy. Accordingly, any proponent of indeterminacy in whatever area must supply an assurance that he has considered *all* the relevant evidence. The relativist in the present debate certainly has not done so.

3. Inter- and intra-cultural interpretation

3.1 In earlier chapters I stressed the conceptual importance of the relation between interpretation and expression. This suggests that the crucial determinant of the issue of universality will be evidence which tends to establish not correlations between behaviour and emotions

in different cultures but whether or not a particular type of behaviour is interpreted in the same way within all cultures and between them. Both Ekman ((1973) pp. 191f) and Birdwhistell ((1971) p. 31, quoted below) agree on the importance of this aspect of the problem. There are two ways in which a consideration of interpretation across cultures may support or fail to support the statement that ϕ -ing is universally an expression of F-ness.

(a) If it can be established that for some human population, P, members of P interpret each others' ϕ -ings as expressions of F-ness, then this will provide substantial (though strictly incomplete) support for the claim that ϕ -ing is an expression of F-ness in P. If the same premiss can be established for all or most other human populations, then the claim that ϕ -ing is universally an expression of F-ness is more or less justified. If, on the other hand, the way ϕ -ings are interpreted in P does not correspond to the way they are interpreted in any other population, then relativism is conclusively justified insofar as it is reducible to the denial of universalism.

(b) If it can be established that for any two human populations, P and P', between which pedagogic transmission of sufficient intensity has not taken place, members of P and P' interpret ϕ -ings as expressions of F-ness then this entails universalism with respect to ϕ -ing as an expression of F-ness.

However, it is difficult to see how either sort of information could be accessible to an investigator who did not already possess considerable knowledge of the beliefs, values and language of the cultural groups being studied. If, therefore, we treat evidence relating to interpretation as fundamental to the question of universality with respect to expressive behaviour we will be forced to abandon the earlier proposal to accord absolute priority to physiognomic expression in the order of ethnological discovery. It is obvious why this should be so. Lacking any other information about the cultural group in question, the only clue to their interpretation of expressive behaviour will be the behaviour it elicits from them in response. The simplest case will be that of reciprocation. But suppose it is found that members of P ϕ in response to ϕ -ings by other members of P and the same holds for P'. This would be compatible with statements like the following : members of P express hostility by ϕ -ing and they tend to reciprocate with equal hostility while members of P' express friendliness by ϕ -ing and they tend to reciprocate each others'

friendliness in kind.

Inter-cultural interpretation is no better off in this respect than intra-cultural interpretation. Consider the following anecdotal example.⁸ Visitors to the South-West of France find that if they greet the inhabitants of small villages in the area by smiling, their smiles will be interpreted as expressive of aggression or contempt. Paradoxically, if we assume no linguistic evidence or acquaintance with prevailing attitudes in French rural communities, then the possibility of making this relativistic discovery presupposes the universality of other kinds of physiognomic behaviour. We can rely only on the way in which natives reciprocate when smiled at by visitors. Presumably what indicates to the visitor that the local inhabitants construe his smiling under an offensive light will be that they respond with an air of surly hostility. Consequently, the relativistic hypothesis about smiling is supported, if at all, by evidence which presupposes the universality, ie. the universal or inter-cultural interpretability, of physiognomic expressions of hostility.

3.2 It is unnecessary, however, to insist on the exclusion of linguistic evidence at this point in the argument. Verbal testimonies supplied by subjects in response to photographs or sketches presented to them, or their commentary on the behaviour of their neighbours may count for or against cultural relativity and thus retrospectively bear upon the way in which the behaviour of the same subjects initially came to be understood and prescriptively upon the methods to be adopted in future investigations of totally alien cultures.⁹ Thus the relevance of Birdwhistell's observations of variation within the United States, a homogenous *linguistic* community, may be conceded.

In one part of the country, an unsmiling individual might be queried as to whether he was "angry about something", while in another, the smiling individual might be asked, "What's funny?". (Birdwhistell (op cit.) p. 31)

Observations of this kind, however, offer only inconclusive support to relativism. There are several psychological variables to be taken into account which complicate matters considerably. The complexities thus

8. Suggested to me by Mark Sainsbury.

9. Birdwhistell finds the evidence supplied by such methods indeterminate (except where it seems to support the relativism to which he is committed). See Birdwhistell ((op cit.) p. 32). For a different view, see eg. Ekman ((1973) pp. 191f), Ekman, Friesen and Ellsworth ((1972) chapters XIII-XIX), Eible Eibesfeldt (op cit.).

introduced may be appealed to by the universalist at least in defending the enduring *possibility* of his theory against the more facile arguments of the relativists (exemplified in an extreme form by LaBarre (op cit.)). In some cases the attempt to make sense of these complexities will sometimes seem to give positive support to the universalist case.

In this connection it will be convenient to revert to the notion of emotional taxonomy mentioned earlier. The classification of emotions is to some extent apriori and certain disagreements among anthropologists can be traced to diverging intuitions about these classifications. Thus Birdwhistell (loc cit.) criticises a universalist strategy which may be adopted to cope with observations of the kind quoted, on the ground that it appeals to an implausibly 'elastic' conception of pleasure (cp. Eible Eibesfeldt's general 'Yes to social contact' quoted earlier) capable of being associated with such divergent attitudes as humour, ridicule, friendliness, hostility, etc. To some extent, as noted in Chapter 6, controversies of this order cannot be settled with any confidence. However, the issue is not simply a matter of how far particular emotional concepts can be stretched before they are deprived of meaning altogether. Clearly there can be no satisfactory answer to questions about which emotions are species of which. But there are apriori links between different emotional concepts which cannot be ignored. These connections are not merely taxonomic, but explanatory. Let us take a rudimentary example. The concept of pleasure is essentially bound up with the notion of gratification and there plainly are indefinitely many kinds of gratification. Two plausible apriori schemata involving pleasure, some other attitude and a kind of eliciting situation are the following. Hostility involves a tendency to feel pleasure at the spectacle of misfortune befalling the object of hostility. Friendliness involves a tendency to feel pleasure at the advantages accruing to the friend. Accordingly, observations to the effect that smiling is sometimes associated with hostility, sometimes with friendliness, that smiles sometimes accompany apologies and sometimes insults, have no bearing on the question of whether or not smiling is a universal expression of pleasure. The same applies to variations in interpretations of smiling by members of different cultural groups which may well be determined by which of the possible apriori schemes is stressed in the group and the hypotheses members of the group tend to form in advance about the attitudes of strangers.

3.3 The failure to acknowledge the conceptual relations between different emotions, or emotions and attitudes, is connected with a tendency to neglect differences between interpretation (or understanding) and the formation of explanatory hypotheses, between first and second order interpretation, and even between interpretation and moral evaluation. The first and second mentioned differences are not themselves sharply distinguishable in practice. This is not important, however, since it is often clear that the apparent plausibility of a conclusion is based on neglect of one of them, even when it is not clear which. The distinction between first and second order interpretation is most easily grasped in relation to language. Let us assume that the first order interpretation of a language, L, is successful when we have a theory of truth for the sentences of L. The truth theory includes a standard theorem of the form : $T(\sigma) \leftrightarrow p$. But it turns out that L-speakers sometimes utter tokens of σ in order to communicate that q. Furthermore, we have reason to believe the following counterfactual : If it were not the case that $T(\sigma) \leftrightarrow p$ in L, then L-speakers would be unable to communicate that q by uttering tokens of σ . If these conditions hold, and if they capture our reasons for interpreting some token utterances of σ as acts of communicating that q, then this interpretation is second order in the sense that it presupposes the first order interpretation which yielded the truth theorem for the *type*, σ . Of course this is a highly abstract characterisation of second order interpretation; it is intended to accommodate the interpretation of various non-literal speech-acts involving metaphor, irony, various kinds of conversational implicature and so on.

There is every reason to suppose that an analogous distinction can be made out for non-linguistic expressive behaviour. The only constraint on the analogy is that noted in Chapter 6, §2.1 where the possibility was admitted that $T(\sigma) \leftrightarrow p$ could be true for L on structural grounds even though L-speakers always used σ to communicate that q, never to communicate that p. The universalist, then, could appeal to the distinction in order to justify the following possibility, ignored by relativists. Members of all human communities recognise a first order interpretation of ϕ -ing as expressive of F-ness, but members of one group, P, recognise a second order interpretation of ϕ -ing as expressive of F'-ness while members of another group, P', recognise a second order interpretation of ϕ -ing as

expressive of F*-ness. 'Fx' does not entail 'F'x' or 'F*x' (or 'F'x or F*x'). Furthermore, 'F'x' and 'F*x' cannot both be true. But, ϕ -tokens are interpretable as expressive either of F'-ness or of F*-ness only in virtue of a shared antecedent interpretation of ϕ -ing as expressive of F-ness.

For example, a certain facial configuration may be taken by us as an expression of easy-going bonhomie. When we try out the same physiognomic expression on members of an alien community, they take it as expressive of disrespect. Birdwhistell, Leach (1972) and others will often argue for relativistic conclusions on the basis of just this kind of discrepancy. But *prima facie* it is just as plausible to maintain that members of the alien community recognise the facial configuration as expressive of relaxed bonhomie and regard such behaviour in a stranger as a slur on their dignity.

Further indeterminacies are introduced by the distinctions between interpretation and explanation and moral evaluation. Thus, should an ethnologist laugh in the presence of a native spectator, the latter may believe the ethnologist to be friendly or he may believe him to be hostile. Both explanations are compatible with a native interpretation of laughter as expressive of amusement, but laughing *at* someone may be a result of a basically hostile attitude, while laughing with someone suggests solidarity. Whichever explanation the native thinks is true, it will be based in a first order interpretation which he shares with the ethnologist.¹⁰

Exactly the same argument can be used to block inferences from discrepancies in evaluation to discrepancies in interpretation. Again, evaluative discrepancies often presuppose constancy of interpretation. Where two audiences recognise the "message" conveyed, one may disapprove either of the "message" itself or of the particular medium adopted while the other accepts the behaviour in question without moral judgement. The discrepancy in evaluation would be a result of discrepancies in the display rules acknowledged by the

10. A striking example of this kind of confusion is exhibited by Leach ((op cit.) p. 324). Leach contrasts the different views of Catholicism held by a Catholic priest and a psychoanalyst, among others. But the psychoanalyst is interested in *explaining* belief in Roman Catholicism in terms, perhaps, of mass neurosis. To do this he must first establish what those beliefs are which he can only do by *interpreting* Catholic dogma and ritual in the same way as Catholics themselves.

different interpreters. ¹¹

3.4 Finally, before proceeding to the last, positive phase of the argument, a particularly disingenuous manoeuvre, common among cultural relativists, should be noted. The basic argument of the relativist runs as follows :

- (i) ϕ -ing is an expression of F-ness in P
- (ii) ϕ -ing is an expression of F'-ness in P'
- \therefore ϕ -ing is not universally an expression either of F-ness or of F'-ness.

But in many cases this form of argument represents as *inter-cultural* variations which are in fact *intra-cultural*. ϕ -ing may be an expression of F-ness *and* F'-ness in all of the cultural groups compared. This fallacy is implicit in the passage quoted from Birdwhistell at the beginning of this section. If his observation is to have any weight, he must also show that in some groups smiling is *never* an expression of amusement. The same objection applies to LaBarre's (op cit.) observation that in some cultures tongue-protrusions are regarded as erotic gestures while in Western European cultures they are considered to be slightly offensive ('mischievous' might be a better description). As an argument for relativism with respect to the expressive properties of tongue-protrusions this fails for two reasons. First, it is vitiated by an intolerable vagueness of description : erotic tongue-protrusions are *morphologically* distinct from mischievous ones. Secondly, within Western Europe, tongue-protrusion is sometimes erotic and sometimes mischievous.

Anticipating this kind of criticism, Birdwhistell ((op cit.) p. 33) finds himself forced to concede that there is no evidence of cultural communities whose members never smile in situations which we would regard as appropriate elicitors (cp. Leach (op cit.) p. 339). But if there is no such evidence, then there is no evidence for relativism as opposed to universalism. The latter theory asserts just that some forms of expressive behaviour are universal.

11. Leach again supplies an illustration of this kind of confusion. At op cit. p. 319 he offers as an example of the 'context-dependency' of conventional symbols the case of a woman dressed and adorned in such a way as to look smart to some and ridiculous to others.

4. The innateness of physiognomy

4.1 Ethnology cannot provide evidence for or against theories of the ontogeny of behaviour. This is so partly for reasons adverted to in passing at the beginning of this chapter. ϕ -ing may be universally an expression of F-ness and yet be learnt because of the possibility of pedagogic transmission between cultures or common environmental features relevant to learning (though it is unclear what the latter could be in the case of expressive behaviour). Again, in principle ϕ -ing might be expressive of F-ness only in one or two cultures and yet be transmitted through inheritance if those cultures were in some sense "racially" distinct. Admittedly, neither of these suggestions is very plausible. Universality would strongly suggest that the behaviour pattern in question is innate given the difficulty of specifying environmental factors which do not themselves include the expressive behaviour of adults and the improbability of sufficiently widespread pedagogic transmission. Much more damaging to the attempt to argue from ethnological premisses to ontogenetic conclusions is the impossibility of determining the truth-value of the said premisses by means of ethnological comparisons alone. I propose now to consider the prospects for a move in the opposite direction, ie. from ontogenetic premisses established by clinical methods, to ethnological conclusions. It seems to me that this procedure is more promising. First, if it can be established that a behaviour pattern is innate - ie. not learnt - then there is good reason to expect it to be repeated throughout the species where this cannot be decided by other means. Likewise, if the behaviour is found to be learnt then we would not expect it to be universal. Even if these inferences are doubted - they plainly are not fool-proof - the issue which primarily interests me is ontogenetic, and this, in the light of the previous arguments, is not best approached via ethnology. In the area of developmental psychology there is, of course, evidential under-determination but not to the same degree as in ethnology. I will argue that in the light of the available evidence there is at least a persuasive hypothetico-deductive argument for innateness.

4.2 The strategy followed here will be analogous to some extent to that adopted for the appraisal of a classical empiricist approach to the ethnology of behaviour. The analogy to be pursued is not between the attitudes of the ethnologist and that of the developmental psychologist, but between the attitude of the former and the primitive

cognitive condition of the infant on entering the world. Just as the empiricist who enters the field of ethnological investigation entirely without assumptions about physiognomic behaviour cannot begin to verify one or other of the theories considered previously, so the infant whose mind is 'a perfect and absolute blank' cannot begin to learn how to behave or how to interpret the behaviour of others. Indeed, the empiricist aspires to a condition of innocence, as conceived by the behaviourism of Watson and his followers, free from preconceptions and predispositions. The methodology of the empiricist reflects the theory of the behaviourist. Against this, I will argue that the infant must be equipped from the outset with a capacity for correct physiognomic perception and appropriate expressive behaviour. Without these initial advantages, it would be, if not impossible, at least extremely difficult to learn how to interpret correctly and perform appropriately. It is worth reiterating, however, that no *conceptual* abilities are imputed to the neonate .

The psychological behaviourist, then, requires the infant to discover his own physiognomic repertoire and the significance of the physiognomies in his environment. How might he achieve this, and what evidence would determine whether or not he had done so? Starting with performance it might be suggested that the infant learns to exhibit expressive behaviour in appropriate situations by imitating the behaviour of adults in situations of that kind. In order to do this the infant must learn to distinguish appropriate from inappropriate situations. It is not clear how this would be achieved. One suggestion would be that the infant learns to exhibit appropriate expressive behaviour in those situations which stimulate in him the emotion expressed. But there is no guarantee that a given situation will induce similar emotions in the infant and in the adults he is supposed to imitate. A second problem is this. How is the infant to appraise the success of his attempts at imitation? Presumably some kind of reinforcement mechanism has to come into play. However, there is no evidence that physiognomic failure invites punishment or that physiognomic success invites reward.¹² Rather, the infant would have to

12. Admittedly crying is sometimes regarded as a device for attracting sympathetic attention. In this way it is actually rewarded, but crying offers no support to the behaviourist. It occurs in most cases immediately after birth and regularly thereafter (see Charlesworth and Kreutzer's essay in Ekman (1973)). It is simply false that the infant is obliged to experiment with various noises and gestures until he discovers one which will produce the desired effect. He gets it right from the start.

rely on adult expression of approval and disapproval, so he must first of all learn to recognise these.

If the learning theory is to be defended, therefore, an account must be given of how the infant first acquires the ability to interpret the behaviour of those around him. But the claim that this involves some kind of learning process faces parallel difficulties. As a first shot it might be argued that the infant is capable of recognising certain situations as elicitors of, say, pleasure. He then "observes" the way adults behave in those situations and "infers" that the observed behaviour expresses pleasure. The infant's primary recognition of the relevant E-situations is determined, presumably, by the emotions induced in him by those situations. But if that is so, and it is difficult to see what alternative basis there could be for infantile discrimination of E-situations, then the theory is again vitiated by the absence of any guarantee that there are any situations which elicit pleasure in both adults and infants. It is, therefore, highly unlikely that the infant learns to understand adult behaviour by correlating it with the observable situations in which it occurs in the manner of a radical translator or empirical ethnologist.

4.3 It is open to the behaviourist at this stage to object that I have represented the learning situation in an unnecessarily complex way. The infant does not need to correlate adult behaviour with E-situations and then imitate it. Rather, if the infant proceeds by imitating adult behaviour as it occurs he will already be exhibiting appropriate expressive behaviour. In other words, we should focus on the question of how the infant learns to *reciprocate* adult expressive behaviour and it is plausible to maintain that he does so by imitating it.¹³

However, there is a considerable gap between imitation and reciprocation. First, reciprocation is not always a matter of reproducing emotionally similar behaviour, but in many cases involves producing complementary behaviour. The appropriate response to aggression is either mutual aggression or it is fear. The evidence of clinical psychology seems to suggest that the ability to express fear in response to aggression is manifested at a very early stage (see Charlesworth and Kreutzer (op cit.)). There is no justification for

13. For the contrast between reciprocation and imitation cp. Merleau-Ponty ((1964) p. 40) where he distinguishes authentic from inauthentic imitation.

the claim that the ability is based on imitation; the infant may never have encountered fear-behaviour in others. Furthermore, the aggressive behaviour capable of eliciting fear-responses in infants is itself expressive and it is most unlikely that the infant will have had the opportunity to discern a correlation between aggressive behaviour and actual danger or damage to himself.

Even if we confine our attention to simpler cases of a behaviour pattern, ϕ , which is appropriately reciprocated by ϕ -ing, the learning theorist's account is highly implausible. Among clinical psychologists there appears to be widespread agreement that infants begin to exhibit 'social smiling' between the fourth and sixth months (ibid.). In other words, within the first six months of his life the infant has begun to respond to adult smiles by smiling. The innatism defended here explains this as reciprocation rather than imitation. According to this view the infant recognises the adult smile as an expression of friendliness or affection and responds by expressing his own pleasure or mutual affection. The main ground for this view is the monstrously developed acumen imputed to the infant by the behaviourist alternative. In order to imitate the behaviour he encounters the infant must be in a position to compare it to his own behaviour. However, particularly in the case of facial expression, our own performances and those of others are experienced in entirely different ways. The proprioceptive or 'introceptive' sensations associated with the motor activity of our own faces and the visual or 'extroceptive' experiences we have of the facial movements of others are incomparable.

It is worth noting in this connection an important, albeit self-evident, contrast between physiognomic and vocal behaviour. We experience vocal, but not physiognomic, performances by ourselves and by others via the same sense modality. This, combined with the fact that even the most rudimentary linguistic behaviour is not manifested until a considerable time after an extensive portion of the adult physiognomic repertoire has already been mastered, underlies the decisive plausibility of the claim that physiognomic behaviour is, in contrast to language, largely innate. Ordinarily the same linguistic skill is exhibited in understanding and performance. Indeed, it is plausible to attribute priority in the explanatory order to the former. The infant begins to produce correct linguistic utterances on the basis of the incipient understanding he has achieved of the linguistic behaviour of his seniors. But, for the reasons given, we cannot, in

the case of physiognomy, appeal to interpretational competence to explain performative competence or vice-versa (cf. Chapter 6, §3.4). Yet despite the absence of a potential explanatory link between the acquisition of performative and interpretational physiognomic skills, their manifestation in reciprocation is simultaneous in principle. Again there is a contrast between physiognomic expression or communication and language; it is not impossible, or even unheard of, for interpretational competence with fragments of the latter to be acquired and manifested independently of and prior to the ability to utter the same sentences. (The reverse is also conceivable.) Obvious examples of this include infantile obedience to commands exhibited before anything recognisable as a linguistic utterance has passed the infant's lips, and the possession of purely "passive" competence in a foreign tongue, professions of which are entirely commonplace.

At the risk of tedium, a caveat, by now familiar, needs to be reiterated. The innatism defended here does not involve attributing to the infant anything that would normally be identified as knowledge. I am not claiming either that the infant *knows* of the smiling adults in his perceptual environment that they are expressing happiness or that the reciprocating infant is aware that he himself is smiling. Such claims are certainly implausible, but the innatism defended here consists partly in the claim that no knowledge is required for there to be a disposition to smile under certain circumstances or a capacity for physiognomic perception. There is in this context, no justification for the view that we are non-inferentially aware that we are smiling. Admittedly there is available a kind of non-inferential awareness that our faces are moving. But to describe that movement as a smile is to say something about how it looks to others, not about how it feels to the agent. Nothing in the way of innate "knowing that" is attributed to the infant to explain his inborn ability and disposition to smile. Similarly, knowledge-that is not ascribed to the infant to explain his capacity for physiognomic perception and reciprocation. With respect to the former, the phenomenological claims defended earlier and the ontogenetic account being put forward here reinforce each other. Perception is intuitively thought to be a more primitive cognitive process than thought. To the extent that physiognomic perception resembles ordinary perception any intuitive resistance to the claim that the former is innate will be weakened. In order to accommodate physiognomic perception and reciprocation a biological theory might postulate the existence of phylogenetically fixed "receptor mechanisms"

and "releasing mechanisms" respectively (Eible-Eibesfeldt (op cit.) p. 307). In less "scientific" terms this comes out as the claim that just as adult displays of aggression induce fear and its manifestation in the infant, so adult displays of affection attract a mutually affectionate response.

4.4 These reflections suggest a rather novel way of staging the conflict between innatism and learning theory with respect to physiognomic expression. For, contrary, perhaps, to some expectations it transpires that the learning theorist, though committed to a *tabula rasa* theory of the neonate mind, must be prepared to attribute to the infant of four months or so far more knowledge than the innatist needs to attribute to him. Reciprocation does not require knowledge in a full-blooded sense : imitation does. In order to estimate our success in imitation we must have access of some kind to the way we look. Admittedly the issue of whether infantile physiognomic behaviour is to be explained on the basis of imitation or of reciprocation is not fully determined by any conceivable evidence (except in the case of behaviour, such as crying, exhibited virtually at the moment of birth). Thus, however implausible in principle, it is logically possible that the infant who smiles at the age of four months is imitating the behaviour of adults who smile at him.

However, there is indirect evidence which tells forcefully against such a claim. First, it seems that many motor activities which are not apparently more complex than facial expression, are not mastered until much later and are sometimes not mastered at all by children who are disabled in one way or another but who do, nevertheless, exhibit appropriate physiognomic behaviour (see Eible-Eibesfeldt (op cit.) p. 304). Secondly, and more pertinently, it appears that the acquisition of a 'specular image' (Lacan (1949)), postdates the exhibition of appropriate physiognomic behaviour by several months, if not years (see Merleau-Ponty (1960)). By 'specular image' in this context, I mean the conception of oneself as an entity possessing a physiognomic surface accessible to visual observation by others (I do not mean a precise internal image of oneself as qualitatively distinct from others which could only be obtained through portraits or reflected images). Phases in the development of the specular image are manifested in the individual's response to his own image reflected in a mirror. Roughly, there is an initial phase through which many animals pass in which the mirror image is regarded as a separate

individual. According to Lacan (*ibid.*), pigeons never advance beyond this phase. An intermediate phase is marked by apparent indifference to the mirror image once it is discovered not to be another animal. The development of most higher mammals, including primates, is arrested at this stage. In human children, however, these phases are followed by a third in which interest in the mirror image is restored. The explanation put forward by Merleau-Ponty (*ibid.*) is, in a highly compressed form, that humans acquire a specular image which enables them to perceive the mirror image as an image of themselves.

Strictly this is inconclusive since it may be that all mammals acquire specular images but only humans are sufficiently vain to become interested as a result in how they look. Nevertheless the evidence is suggestive and its relevance here is plain. For the preceding criticism of learning theory can be recast thus : in order to imitate others successfully the infant must first acquire a specular image. But the evidence alluded to just now indicates beyond reasonable (ie. nonphilosophical) doubt that the specular image is not acquired until long after the human physiognomic repertoire has become operative. Consequently the elements of that repertoire cannot be obtained through imitation and it is doubtful whether any alternative learning or conditioning theory is available.

4.5 The arguments of §§4.3 and 4.4, relying as they do on apriori assumptions about when an individual can safely be regarded as capable of learning various cognitive skills and on the support of indirect evidence, are best summarised in a hypothetico-deductive form. This will be in keeping with the procedures adopted by innatists from Darwin to Ekman; the latter ((1973) p. 6) quotes the former as follows :

No one could be a good observer unless he was an active theoriser. In the light of the impotence of unaided induction to settle the issues discussed in this chapter and given the respectability of deductive modes of inference even in modern physics, it is surely legitimate to argue that among the logically possible theories of the provenance of physiognomic competence which are compatible with the data delivered by observation of infantile behaviour, the theory of innate capacities for physiognomic performance and perception is the best available. Of course, this claim needs to be fleshed out with some mention of the relevant criteria by which one theory is judged to be better than another in advance of its trial by observation and a justification of

the assurance, implicit in hypothetico-deductive techniques, that all the competing theories have been considered *in senso composito* if not *in sensu diviso*. Fortunately, these requirements are fairly easy to meet in the present context.

Regarding the first requirement, there are a number of criteria whose general scientific applicability in the absence of empirical counter-indications is widely recognised; these include, notably, simplicity and conservatism. As analogues in developmental psychology, the following principles suggest themselves respectively :

(i) With respect to a given type of behaviour tokens of which individuals are manifestly able to perform, assume the lowest degree of cognitive development sufficient to generate behaviour of that kind.

(ii) Assume, as far as possible, the same degree of cognitive development to explain all the different kinds of behaviour exhibited by an individual at any given phase of his development.

It is plausible to maintain that as guiding principles, (i) and (ii) capture the most basic motivation of behaviourist learning theories in general. Yet the arguments of the preceding paragraph indicate that innatism with respect to facial expression approaches their satisfaction more nearly than learning theory.

As regards the second question - how can we be certain that all the coherent theories have been considered? - the reply is simply that the field is divided exhaustively by the theories we have considered at the level of generality at which we have considered them. 'Either facial expression is innate or it is learnt' instantiates the law of the excluded middle. Insofar as a disposition or propensity is innate only if it is not learnt there is no conceivable intermediate theory. This is to be understood, naturally, in such a way as to allow for the subsequent intervention of conventions or pathognomic disturbances, conformity to which is a learnt disposition, which may have the effect of inhibiting the manifestation of innate propensities.

4.6 If it is true that there is an innate contribution to the ability to perform and understand some expressive behaviour, and the case for this claim is certainly strong, then how does this help the ethnologist? Probably it does not help him as much as it aids the infantile explorer of his own culture. The main contribution of the ontogenetic theory to ethnology consists in the permission it gives the ethnologist to trust his own physiognomic perception. He may, being a member of the same species as the subjects of his

investigation and, as such, influenced by the same biological forces, rely, to some extent uncritically, upon the imponderable evidence supplied by native physiognomic behaviour. True, he cannot proceed without reservations, for while θ_{iii} above appears to be false for such expressions as smiling, the potential interference of display-rules needs to be anticipated. However, if it is conceded that the investigator knows in advance what some native behaviour means, then he is in a position to draw substantial cultural conclusions from correlations between instances of such behaviour and the E-situations in which they occur. Sometimes, no doubt, he will be unable to determine, even after prolonged acquaintance, whether members of a particular cultural group actually feel what they express or merely express certain emotions in order to conform to some cultural convention. But a restricted indeterminacy of this kind may be insuperable; indeed, even in relation to the behaviour of our compatriots, or our own behaviour, we cannot always give a determinate answer to questions of this kind. Often we are not even disposed to raise the question. It is, in some contexts, both a social solecism and an epistemological mistake to discriminate too sharply between sincerity and insincerity, honesty and hypocrisy, without sufficient sensitivity to the subtle exigencies of courtesy and decorum.

VIII. CONCLUDING NOTE ON THE PHILOSOPHICAL RELEVANCE OF NATURAL
EXPRESSION

1. In Chapter 5 I stressed the impotence of theories of natural expression to supply a refutation of scepticism about other minds. However, in recent philosophical debate considerable emphasis has been placed on problems which occur at certain points of intersection between metaphysics, semantics and the theory of concept-acquisition. I wish to conclude this essay by sketching, in a tentative and programmatic way, a possible solution to one such problem. The problem concerns the acquisition of psychological concepts, the formulation of semantic theories for the psychological parts of natural languages and the metaphysical issue of the reality of the mental. The suggested solution is to exploit the theory of natural expression whose relevance to the acquisition of psychological concepts was noted by Wittgenstein in *The Philosophical Investigations*.

Difficulties are raised for standard truth-conditional semantics by sentences which cannot fully be verified by observation alone. 'Realist' theories of this kind seem to be forced to mention, in their semantic descriptions of some sentences, truth-conditions which include sorts of circumstance inaccessible in principle to observation. If this is actually the case, then such theories fail to meet certain requirements reasonably stressed by proponents of verificationism. Specifically, they cannot explain how the ability to understand the sentences thus defined could possibly be acquired or manifested, or, indeed, how we can even justify the claim that we possess such conceptual abilities ourselves. More accurately realist theories appear to make the acquisition, possession and manifestation of possession of certain concepts impossible (for these points see Dummett, eg. (1959) p.67 : (1973) p.467 : (1976) p. 80f). Among the sentences which are problematic in this way are those which we use to attribute mental states to people other than ourselves. Indeed, for reasons shortly to be made explicit, self-ascriptions are no less problematic than other-ascriptions (cf. Strawson (1959) p. 106). I suggest, however, that there is a form of realism about the mental which, in conjunction with the foregoing, goes some way towards overcoming these difficulties. The foregoing account of natural expression offers some support to a theory proposed by McDowell (1978) which, for brevity, I will label 'immanent realism', thus contrasting it

as McDowell himself does, to the usual, transcendent variety of realism. The question, then, is whether realist theories of mental concepts are bound to be transcendent in a sense to be clarified shortly.

2. The transcendent realist thinks that he can avoid these difficulties by appealing to a 'truth-value link' as the principle underlying competence with other-ascriptions of psychological predicates. According to this kind of realist, we learn the meaning of psychological predicates initially in their application to ourselves. This is assumed to be unproblematic, even by strict verificationist standards, the relevant truth-condition being readily accessible to the learner's consciousness. Then, having acquired a degree of competence in the use of personal pronouns - obviously in other, less problematic contexts - we combine the two skills in order to understand ascriptions to others of psychological predicates we already know how to apply to ourselves. 'He is in pain' is true just if 'I am in pain' uttered by the person to whom the third personal pronoun refers, would also be true (McDowell (1978) p.130). Thus the transcendent realist contends that although another person's being in pain is a circumstance in principle inaccessible to my consciousness, it is, nevertheless, an instance of a kind of circumstance other instances of which are accessible to me, namely when I am the person who is suffering.

While conceding that truth-value links of this kind constitute an ineliminable aspect of the semantics of psychological discourse, McDowell denies that they can contribute anything to the explanation of the mastery of the concepts in question.

This, so far from solving the problem, simply ignores it. If someone cannot see how another person's being in pain - on an interpretation of that particular circumstance which makes it inaccessible - can possibly enter into the meaning one attaches to some form of words, one does not allay his worry by baldly reasserting that it does.' ((ibid.) p.132)

Admittedly, McDowell's presentation of the transcendent realist's case is partly vitiated by a certain oscillation between the sorts of circumstance to be considered and between taking sorts of circumstance and instances as primarily relevant. If the transcendent realist is represented as having made certain decisions on these points and sticking to them, then the quoted criticism loses much of its force. A consistent realist of this kind will deploy two *kinds* of circumstance in his explanation of the mastery of psychological concepts in general. One kind of circumstance, that of *another* person's being in a certain emotional state or having a certain sensation, is such that in principle

instances of it are never accessible to consciousness. The other relevant *kind* of circumstance, that of *someone* being in the same kind of psychological state, is sometimes accessible, namely, when I am in that state. The transcendent realist's case depends crucially on the fact that all instances of the first kind are instances of the second kind of circumstance but not vice-versa.

McDowell, in effect, has slightly mislocated the defect in transcendent realism. Proponents of this doctrine, as portrayed by McDowell, assume as unproblematic the prior ability to apply psychological concepts to ourselves. True, this does not violate the requirement emphasised in the debate so far, which states only that the truth conditions of the relevant sentences must be accessible to the person whose competence with those sentences is to be accounted for. But satisfaction of that requirement alone simply is not sufficient for the explanation of the mastery of psychological concepts even if we confine our attention to their use in the first person. It is not legitimate, therefore, to start theorising from the assumption that we know what pain or distress is from our own case (McDowell (*ibid.*) p.131). The assumption is simply false if it is taken to refer to a phase in our cognitive development which is anterior to the acquisition of linguistic competence. The most that can be conceded is that, unless we are unusually fortunate, we are personally acquainted with distress from an early stage, and are able to recognise repeated instances of it as similar to past experiences. But the discovery that these experiences are instances of distress, the discovery, in other words, that on these occasions the concept of distress is applicable to ourselves, is no more or less problematic than the discovery that, on other occasions, the same concept is to be applied to others.

The transcendent realist, then, has no answer to Wittgenstein's point that the criteria for the use of a term must be public (Wittgenstein (1953) eg. §§ 257, 579). This is to say that the conditions for the application of psychological vocabulary, whether or not we regard them as truth-conditions, must be such that instances of them are at least sometimes accessible (in principle) to more than one observer simultaneously. Above all, this requirement must be met in the paradigm situations which permit the pedagogic transmission of the concepts to take place. In these situations it is crucial not only that the learner should be in a position to tell when it is correct to apply psychological terms to others, but also that the teacher should be in a position to tell when it is correct for the learner to apply those terms to himself.

If the emotional states of others are in principle inaccessible, then competence with self-ascription of emotion could never be manifested, thus failing to meet two of Dummett's demands simultaneously. Self-ascriptions are no less problematic than other-ascriptions of psychological terms since, on the transcendent view, there is nothing accessible to the public gaze which will allow the teacher or the learner, or, indeed, anybody, to ascertain whether or not psychological terms are being correctly self-applied. The teacher cannot correct the learner's misapplications to himself of the problematic concepts, or reinforce his proper applications of those concepts to himself. Conversely, the learner cannot see why others use those concepts when they do so, and others will have no basis for applying them to the learner.

3. Immanent realism purports to solve these problems, and thereby preempt anti-realism, by endowing physiognomic perception with a serious role in the explanation of the mastery of psychological concepts. McDowell introduces this refurbished realism, which I believe to be vindicated by considerations advanced in Chapters 5 and 7, as follows :

In the view of this different realist, then, we should not jib at, or interpret away, the common sense thought that, on those occasions which are paradigmatically suitable for training in the assertoric use of the relevant part of a language, one can literally perceive, in another person's facial expression or his behaviour, that he is in pain, and not just infer that he is in pain from what one perceives.' (McDowell (1978) p.136)

Little needs to be added to what has already been said in Chapter 5 about the concept of physiognomic perception upon which immanent realism with respect to the emotions of others is based. Of course, the concept of immanence, as it occurs in the label, is not to be confused with the concept of perceptual immediacy considered in Chapter 5. All that immanent realism requires is that the kinds of objects to which it applies should sometimes be accessible to observation unaided by inference. Physiognomic perception permits us to make this claim about the affectivity of others. Realism based on physiognomic perception cannot be accused of demanding too much active intellectual participation from the infantile language learner (cf. McDowell (ibid.) p.140 where references to passages in Wittgenstein are also supplied). There are at this stage two available means of meeting objections of this kind and thus diminishing their tendency to support anti-realism. First, the infant relies in the relevant paradigm contexts on a kind of perception which is not sharply distinct from the perceptual abilities upon which he relies in coming to grips with purely material concepts. Secondly, the infant does indeed make some kind of active contribution to the learning process. In the

acquisition of ordinary, non-psychological concepts, he is expected to make certain observations and to exploit certain primitive capacities for visual recognition and discrimination. The acquisition of psychological concepts also involves the operation of certain cognitive devices; devices with which it seems we are innately endowed (cf. Chapter 7).

It would, however, be disingenuous to insist that immanent realism satisfies precisely the criteria stressed by anti-realists in the construction of theories of meaning. To begin with, no realist can consistently attempt to respond to a need sometimes alleged to admit into the theory of meaning only sorts of conditions, instances of whose satisfaction can always be observed to hold. If anti-realism is based on the premiss that competence must be *exhaustively* manifested, ie. unequivocally disclosed on every occasion of its exercise, then it imposes a requirement which is implausibly stringent. I have nothing to add to McDowell's own thoughts on this issue (*ibid.* §7). A more central criticism would be the following: immanent realism, so-called, fails to provide a genuine response to the requirement that a theory of meaning should be genuinely empirical. It merely replaces that requirement insofar as it is imposed on the acquisition and manifestation of concept-mastery, with a different, *phenomenological* requirement.

This is true, but there is no justification for regarding this as a relaxation of the methodological principles observed by orthodox empiricists, as a movement away from a certain standard of rigour. A more accurate perspective is obtained by casting our thought in the spirit of Max Scheler, according to whom there is a certain sense in which

... phenomenological philosophy is the most radical *empiricism* and positivism. (Scheler (1913) p.138 in (1973))

The peculiar rigour of the phenomenologist consists in a refusal to rely upon extraneous criteria of experience, preferring to focus upon what he calls lived experience itself. This involves an alternative but, in my opinion, authentic kind of investigative stringency exemplified not only (or even especially) by those writing within the "phenomenological movement", but also by those who, like Wittgenstein and Wollheim, refuse to take for granted traditional assumptions about the scope of observability. Admittedly, as McDowell points out about Wittgenstein, the true phenomenologist will refuse 'to take sides in the typically *philosophical* debate between transcendent (truth-value-link) realism and anti-realism' (*ibid.* p.136n). Here, however, the role of

phenomenology is to provide a plausible description of a category of experience as a preliminary to deciding whether a given metaphysical theory, realism, is compatible with it or whether, should a negative answer be confirmed, its antithesis, anti-realism, will be forced on us.

In sharp contrast to the "method" of phenomenology, there is the 'naturalistic' procedure adopted by Quine who relies upon a notion of stimulation to be fixed by criteria involving 'the pattern of chromatic irradiation of the retina' (see Quine (1960) pp.31f). Indeed, the reliability of these criteria is severely impugned by visually ambiguous figures, like the duck-rabbit, which interested Wittgenstein precisely because of their bearing on the concept of perception in general (cf. Føllesdal (1975) p.36 for an intriguing application of these considerations to Quine's method of radical translation); their adequacy is threatened also by a certain kind of illusion, the best known example of which is probably the Muller-Lyer figure.¹ Such illusions can be explained only by admitting the intervention of a distorting influence at a location, so to speak, somewhere between the retinal impact and the receipt of a visual impression. It is surely obvious that where experience and retinal activity diverge, or where the former is inadequately described in purely optical terms, then we must favour the experience itself, rather than an incomplete account of its mechanics, as the basis for the acquisition of conceptual mastery and its manifestation to those who, untrained as they are in the science of optics, are nevertheless the best judges of such matters, namely, native speakers of the language in question.²

Immanent realism, then, will provide an account of the acquisition of psychological concepts which goes some way toward meeting the requirement that meanings or concepts should be firmly anchored to experience, while at the same time preserving for the mental a degree of transcendence or 'privacy' sufficient to do justice to the ordinary preconceptions we have in this area. I think that these are at least good prima-facie reasons for attempting to develop and defend a theory of immanent realism in detail. Unfortunately the task cannot be undertaken here.

1. Other striking examples, involving chromatic as well as morphological distortion, are to be found in Gombrich (1960), illustration VI and Wollheim (1973).

2. Where such divergence is commonplace throughout the species (cf. Chapter 6 § 5.3) the innate mechanisms we postulate in order to explain the provenance of the relevant experiences (as in Chapter 7) should be accorded much the same explanatory status as the purely optical mechanism also involved.

APPENDIX : SPONTANEITY AND PURPOSELESSNESS

1.1 In order to avoid a possible charge of negligence, it will be necessary to bring to the fore a feature, or, more precisely, two features of natural expression hitherto merely alluded to in passing. Thus, at the close of Chapter 3, I indicated that there might be a teleological aspect to the difference between expressive and purposeful action. Again, at Chapter 4 § 2.3, I alluded to the spontaneity of much expressive behaviour which constitutes a ground for scepticism regarding any attempt to elucidate it on the basis of highly complex sets of communicative intentions. The relatively small number of philosophers who have discussed non-verbal expression in any depth includes Alston (1965) and Danto ((1973) Chapter 6). Both were struck by what they took to be the characteristically spontaneous nature of natural expression, but they approached this aspect from different angles. Alston believed expression to be spontaneous in contrast to linguistic communication, allowing, however, for exceptions like interjections (Alston (op cit.) pp.26 f). Danto, on the other hand, introduces a 'principle of spontaneity' intended to contrast expressive behaviour to intentional, or purposive action (no apparent distinction being registered between intentionality and purposiveness, see Danto (op cit.) p. 163). In fact these different routes generate slightly different concepts of spontaneity, and I will refer throughout to the concept disclosed by Danto as 'purposelessness'. There are parallels between the concepts nonetheless.

The reason for not stressing these aspects of expression earlier is that they do not strike me as essential even to the concept of natural expression. Danto's claim to have unearthed a necessary condition of expression in his principle of spontaneity (purposelessness) is, to my mind, profoundly suspect. Admittedly, a certain brand of popular wisdom has it that an inclination to "manipulate" others is incompatible with authentic expression. But from this it would follow at most that purposelessness is necessary for sincerity. To this it might be added that the concept of sincerity is logically bound to the concept of expression in ways somewhat analogous to the links between assertion and truth. However, it seems, *prima facie*, virtually impossible to provide an absolutely watertight distinction between manipulation and

intentions to secure various kinds of uptake characteristic of, though not essential to, communication. It is surely misplaced to accuse one whose passion is genuinely felt, of insincerity in attempting, by expressing it, to obtain its requitement.

Similarly vague intuitions suggest that sincerely expressive behaviour cannot be in any way contrived. But if this is the measure of spontaneity, then it can hardly be appealed to in distinguishing natural expression either from speech or from intentional action in general. My aim in the paragraph which follows is to sketch a more precise notion of spontaneity and of purposelessness, concepts which describe features arguably characteristic of, but by no means universal to expressive behaviour.

1.2 The spontaneity of physiognomic expression can be captured by contrasting it to what are, in a sense, the simplest and most readily investigated kinds of assertion. The relevant utterances are those whose stimuli are accessible to observation by speaker and interpreter simultaneously. At the outset the radical interpreter is bound to rely almost exclusively on the availability of such utterances. However, as Quine pointed out, the presence of the stimulus whose description will eventually enter into the theory of the meaning of the associated utterance type, is not in itself sufficient to elicit a token of it. Further linguistic prompting is also required (Quine (1960) p. 30). For example, with respect to a sentence, σ , of the native's language, such prompting may take the form of interrogation to which, if the appropriate stimulus is present, the native will reply with an assertoric utterance of σ . Alternatively, the interpreter may experiment by uttering tokens of σ in the presence of native speakers and stimulus conditions conjectured to be pertinent, in order to ascertain whether, under those conditions, assertions of σ prompt assent or dissent. In general, utterances of, or assent to σ will be elicited only by a combination of appropriate linguistic prompting and the presence of appropriate perceptual stimuli.

This is not the case for physiognomic expression. Here it seems that typically the presence of an appropriate stimulus is sufficient by itself to elicit the corresponding expressive behaviour. In this respect, the radical interpretation of expressive behaviour across cultures is somewhat simpler than that of linguistic behaviour (cf. Chapter 2 §3.2), thus to some extent confirming the conception of its priority in ethnological research argued for in Chapter 7. Thus, for

example, expressions of (as opposed to accurate descriptions of) distress are typically elicited by whatever causes and is the object of the distress suffered (cp. Danto (op cit.) p. 154). No additional prompting is required. In this sense expressive behaviour is more spontaneous than linguistic assertion.

It might be objected to this that the phenomena of reciprocation, upon which the previous arguments for innateness heavily depend, involve something closely akin to linguistic prompting. Reciprocation might be conceived of as a transaction essentially involving the physiognomic prompting of complementary physiognomic behaviour. However, if this is a potentially correct characterisation of reciprocation, it must be supplemented by indicating that it is a conspicuous feature of reciprocation that no stimulus external to the transaction itself is required. If reciprocation is a matter of prompting and the behaviour prompted, then the prompting behaviour is sufficient for the elicitation of the reciprocal behaviour without the aid of some further stimulus. Accordingly, the typical spontaneity of expressive behaviour consists in its capacity to be elicited either by prompting or by stimulation, but not typically by a combination of both. This is all that is needed to contrast it to the typical elicitation of assertoric observation sentences. Still, the matter could be rather differently construed. The prompting physiognomic behaviour might be thought of as being itself the stimulus of the complementary behaviour it elicits. The choice depends to some extent on whether we describe the innate contribution to the capacity for physiognomic reciprocation in terms of innate releasing mechanisms, or confine the innate material to innate receptor mechanisms. Both conceptions will yield a conception of physiognomic expression as spontaneous in comparison to linguistic communication.

1.3 It would be gratifying, in view of developments in the argument to follow, if we could extract from the foregoing, rather modest conception of spontaneity, a conception of the explanation of expressive behaviour which was, in contrast to the radical interpretation of languages, non-holistic. This would permit the establishment of an elegant parallel between spontaneity and purposelessness, where the latter concept, insofar as it is applicable to naturally expressive behaviour, points to a contrast between it and (intentional) action in general in that the explanation of tokens of the latter, but not of the former, requires the holistic deployment of the concepts of belief and desire. However, the parallel cannot be insisted upon. In the case of radical

interpretation the concepts to be deployed are that of the meaning of an utterance and that of the relevant belief of the speaker (the 'E-concepts', as Peacocke (1979) p. 18 calls them). These concepts might be thought to match stimulus and prompting in that the effect of the stimulus is to produce a certain belief while the effect of prompting through appropriate interrogation is to elicit an expression of that belief in words.

However, the attempt to express the contrast between language and physiognomic expression on the basis of an analogy between belief and emotion respectively, and linguistic meaning and physiognomic meaning (construed, as in Chapter 4, in terms of *prima facie* evidence) respectively, leads to falsehood. Sincere expressions of F-ness involve behaviour exhibited both because the agent is F and because the relevant behaviour constitutes *prima facie* evidence that the agent is F. But this ignores the crucial point that 'because' alludes to different kinds of explanatory project in each case (cf. Chapter 4, p.87 and Chapter 6, §4.2). The spontaneity of expressive behaviour, consequently, can be indicated by contrasting it to certain facts about the explanation of token utterances. If we have a sentence, σ , which means that p in a language, L, then token-utterances of σ by L-speakers are not adequately explained by speakers' believing that p and the meaning of σ . The explanation of an utterance of σ involves the speaker's belief that p and at least the speaker's belief that σ means that p (cf. Chapter 2, pp. 29f). But the agent of a token natural expression of F-ness may have produced that token just because he is F. He need not, and frequently does not, have any further beliefs about what his behaviour expresses. The way in which the latter fact enters into the explanation of such behaviour is not genuinely teleological (cf. Chapter 4, *loc cit.*). To this extent, then, the spontaneity of natural expression is captured by contrasting it to the scheme of holistic explanation associated with radical interpretation.

2. The characteristic purposelessness of tokens of naturally expressive behaviour, noted by Danto (*loc cit.*) and Stuart Hampshire (cf. Chapter 3, p.62), is readily captured by denying the application to such behaviour of the scheme of holistic explanation applicable to intentional action in general. In general, an action of ϕ -ing will not be explained simply by attributing to its agent a desire to bring it about that p without also attributing to him a belief that if he ϕ -s then it will ensue that p (cf. Peacocke, eg. p. 5). Some instances

of expressive behaviour may well be explained by applying a scheme of this kind, but the crucial point is that many will not. For example, a man incarcerated in a firmly wrought cage may express his desire to escape by violently shaking the bars that prevent the realisation of that desire. Our conviction that in such a case what the man is doing is expressing his desire, rather than trying to bring about its satisfaction, depends in part on our belief that he is too intelligent to believe that he will win his freedom by manually destroying the obstacles to it. It is tempting, then, to say simply that ϕ -ing is an expression of F-ness if people tend to ϕ when they are F irrespective of any beliefs they may have about the probable consequences of ϕ -ing. This, however, fails to take into account the possibility of one who desired a certain state of affairs but believed that by expressing that desire he would not merely not contribute to its furtherance, but would actually impede it. The caged man, for example, would refrain from shaking the bars of his cage if he believed that were he to indicate in any way to his gaoler his desire to escape, the gaoler, being of a sadistic disposition, would thereby be encouraged to prolong his imprisonment. Some provision must be made, therefore, for the inhibition of expressive behaviour where that inhibition is itself motivated by beliefs and desires relating to the consequences of expression. Such an allowance is made in Danto's principle of spontaneity (*loc cit.*), and need not in any way impair the characteristic purposelessness of expressive behaviour. All that needs to be admitted is that people will often express their desires and emotions even if they have no beliefs about the consequences of doing so.

Two points need to be made here. Gombrich ((1972) pp. 374f) has pointed out that in many situations like that described in the last paragraph, it will be impossible to determine on the basis of observation alone whether someone is expressing or trying to satisfy a desire or inclination (*cf.* Chapter 3, *loc cit.*). A certain degree of indeterminacy must certainly be accommodated here. But the indeterminacy will be cut down considerably by allowing extended observation of the behaviour over a period, not possible in Gombrich's example of the representation of a calf struggling with its captor, and, as in the example above, assumptions about the agent's intelligence. The criterion of purposelessness might best be formulated counterfactually. x 's ϕ -ing because he desires the realisation of p is an expression of the desire that p if x would have ϕ -ed even if he did not believe that by ϕ -ing he would bring it about that p (so long as he did not believe that ϕ -ing would impede the realisation of p). Given certain assumptions about

intelligence, persistent ϕ -ing on several occasions without results would most plausibly be regarded as expressive. A point that will help to confirm and clarify the conception of purposelessness being promoted here concerns the occurrence of conflicting desires. Desires conflict when their joint realisation is impossible. When one suffers from libidinal conflict of this kind, one cannot, if one recognises the impossibility of satisfying two desires, act with the purpose of satisfying both of them. There is, however, no impediment to the expression of both of them, albeit serially in most cases, not simultaneously.

Secondly, as indicated earlier, purposelessness is not, as Danto thought, a necessary condition of natural expression. People sometimes intentionally express their desires with the purpose of getting others to contribute to their realisation. Admittedly, purposes are not the same as intentions. In one sense people sometimes do things intentionally but for no purpose. Danto, however, thinks that the apparent spontaneity of natural expression entails that tokens of it are, at least characteristically, unintentional (op cit. p. 164). Indeed, there is some connection between purposelessness as I have characterised it, anti-holistically, and lack of intention. For it is, arguably, a necessary condition (not a sufficient one) of an action's being intentional that it should be explained in the right way by some of the agent's beliefs and desires. Thus, for example, smiling may be intentional under the description, 'expressing pleasure', in which case the person smiling must believe that pleasure can be expressed by smiling.

One response to this may be to impose a limit on the degree of purposelessness admitted to be characteristic of expression. Expressive behaviour is behaviour characteristically unaccompanied by any ulterior motive, ie. by any purpose beyond that of expressing some emotion. Danto, however, goes further, and argues that naturally expressive behaviour is behaviour characteristically unaccompanied by any purpose whatsoever, ie. is, at least primarily, unintentional. This puts Danto under some pressure to come up with an account of those cases in which expressive behaviour patently is intentional. His proposal is ingenious and, within limits, plausible. Our naturally expressive behaviour is intentional only in a derivative sense. We learn first to inhibit the expression of, to 'forbear' from expressing the emotions we feel. Having acquired this ability, we are then able to express ourselves intentionally by forbearing from exhibiting the relevant behaviour (elements of a similar idea are discernible in

Hampshire (1960) and Wollheim (1966) p. 102). This works tolerably well for Danto's chosen paradigms, laughter and crying.⁵ However, the principle is not plausibly extended to smiling. The possibility of doing so intentionally is not a matter of double forbearance. For surely smiling is, in Danto's idiom, unproblematically included in our repertoire of basic actions.

I conclude, then, that while spontaneity and purposelessness are, in a loose sense, characteristic but not definitive of expression, the point should not be construed directly in terms of intention. It is sufficient to bear in mind the inapplicability in many cases of the two schemes of holistic explanation. True, those cases in which neither scheme is applicable, will be cases of unintentional expression, spontaneous and purposeless. In some cases, however, beliefs about what our behaviour expresses will have a role in its explanation. These cases are correspondingly less spontaneous, though they may still be purposeless. The crucial point is that our expressive behaviour often is not explicable by reference to such beliefs or to ulterior purposes. In this large class of cases, the explanation of expressive behaviour is non-holistic.

5. Danto does not apply the principle to his third major example, tumescence of the male organ as an "expression" of sexual desire (op cit. pp. 151f), presumably because genital movements are not in any sense directly subject to voluntary control. My own intuition is that erections should not be counted as expressive. However, the temptation to explain this precisely on the basis of the absence of voluntary control should be avoided. Otherwise it would be impossible to accommodate blushing as an expression of shame or embarrassment. What excludes tumescence is, rather, its failure to conform to the functional clause in the definition of expression proposed in Chapter 4. The function of tumescence in humans is procreation, not display. In those primates for whom genital display has an expressive function, this is unrelated to mating; it seems generally to be an expression of aggression.

BIBLIOGRAPHY.

- Aldrich, Virgil (1978), '"Expresses" and "Expressive"', Journal of Aesthetics and Art Criticism, XXXVII, 203-17.
- Alston, William S. (1965), 'Expressing' in Philosophy in America, ed. Max Black (London: Allen and Unwin).
- Anscombe, G.E.M. (1957), Intention (Oxford: Blackwell).
- Argyle, Michael (1975), 'The Syntaxes of Bodily Communication' in Benthall and Polhemus (1975).
- Beardsley, Monroe C. (1971), 'Musical Expression' in Hospers (1971).
- Bennett, Jonathan (1976), Linguistic Behaviour (Cambridge University Press).
- Benthall, J. and Polhemus, Ted (eds.) (1975), The Body as a Medium of Expression (London: Allen Lane).
- Bergson, Henri (1911), Laughter (translation) (London: Macmillan).
- Birdwhistell, Roy (1971), Kinesics and Context (London: Allen Lane).
- Bouwsma, O.K. (1971), 'The Expression Theory of Art' in Hospers (1971).
- Chihara, C.S. and Fodor, J.A. (1965), 'Operationalism and Ordinary Language' in Pitcher (1966).
- Chomsky, N. (1975), 'Knowledge of Language' in Minnesota Studies in the Philosophy of Science, vol. VIII, eds. Keith Gunderson and Maxwell Grover (1975) (Minneapolis: University of Minnesota Press).
- Chomsky, N. (1976), Reflections on Language (New York, London: Temple Smith, Fontana).
- Collingwood, R.G. (1938), The Principles of Art (Oxford: Clarendon Press).
- Cooke, J.W. (1969), 'Human Beings' in Studies in the Philosophy of Wittgenstein ed. P. Winch (1969) (London: Routledge and Kegan Paul).
- Danto, Arthur (1965), 'Basic Actions', American Philosophical Quarterly, II.
- Danto, Arthur (1973), Analytical Philosophy of Action (Cambridge University Press).
- Darwin, Charles (1872), The Expression of the Emotions in Man and Animals, republished 1965 (London and Chicago: University of Chicago Press).
- Davidson, Donald (1963), 'Actions, Reasons and Causes', Journal of Philosophy, LX, 685-700 (page refs. to Davidson 1980).
- Davidson, Donald (1967), 'The Logical Form of Action Sentences' in The Logic of Decision and Action, ed. N. Rescher (Dordrecht: D. Reidel).
- Davidson, Donald (1970), 'Mental Events', in Experience and Theory, eds. Lawrence Foster and J.W. Swanson (Amhurst: University of

- Massachusetts Press). Page refs. to Davidson (1980).
- Davidson, Donald (1971), 'Agency', in Agent, Action and Reason, ed. R. Binkley et. al. (Toronto University Press).
- Davidson, Donald (1973), 'Radical Interpretation', Dialectica, 27, 313-27.
- Davidson, Donald (1974), 'Belief and the Basis of Meaning', Synthese, 27, 309-23.
- Davidson, Donald (1975), 'Thought and Talk' in Guttenplan (ed.) (1975).
- Davidson, Donald (1980), Actions and Events (Oxford: Clarendon Press).
- Dennett, Daniel (1976), 'Conditions of Personhood' in The Identities of Persons, ed. A.O. Rorty (University of California Press).
- Donagan, A. (1966), 'Wittgenstein on Sensation: Cartesianism vs. Behaviourism' in Pitcher (1966).
- Douglas, Mary (1971), 'Do Dogs Laugh? A cross-cultural approach to body symbolism' in Journal of Psychosomatic Research, 15, 387-90. Page refs. to Polhemus (1978).
- Douglas, Mary (ed.) (1973), Rules and Meanings (Harmondsworth: Penguin).
- Dummett, M. (1959), 'Truth', Proceedings of the Aristotelian Society, 59, 141-62. Reprinted in Strawson (ed.) (1967).
- Dummett, M. (1973), Frege: Philosophy of Language (Duckworth).
- Dummett, M. (1976), 'What is a Theory of Meaning? II' in Evans and McDowell (1976).
- Efron, D. (1941), Gesture, Race and Culture (The Hague: Mouton (1972)).
- Eible-Eibesfeldt, I. (1972), 'Similarities and differences between cultures in expressive movements' in Hinde (1972).
- Ekman, Paul (1973), 'Darwin and Cross-Cultural Studies of Facial Expression' in Ekman (ed.) (1973).
- Ekman, Paul (ed.) (1973), Darwin and Facial Expression: A Century of Research in Review (New York: Academic Press).
- Ekman, Paul (1980), 'Biological and Cultural Contributions to Body and Facial Movements in the Expression of the Emotions' in Explaining Emotions, ed. A.O. Rorty (University of California Press).
- Ekman, Paul, Friesen, Wallace and Ellsworth, Phoebe (1972), Emotion in the Human Face (New York: Pergamon Press).
- Evans, Gareth and McDowell, J. (eds.) (1976), Truth and Meaning: Essays in Semantics (Oxford: Clarendon Press).
- Føllesdal, D. (1975), 'Meaning and Experience' in Guttenplan (ed.) (1975).
- Freud, S. (1905), Jokes and Their Relation to the Unconscious, (translation (1960)) (London: Routledge and Kegan Paul).
- Gardiner, P. and Hare, R.M. (1964), 'Pain and Evil', Proceedings of the Aristotelian Society, Supp. Vol. 1964.

- Goldman, A.I. (1970), A Theory of Human Action (Englewood Cliffs, N.J.: Prentice Hall).
- Gombrich, E.H. (1960), Art and Illusion (London: Phaidon).
- Gombrich, E.H. (1963), Meditations on a Hobby Horse (London: Phaidon).
- Gombrich, E.H. (1972), 'Action and Expression in Western Art' in Hinde (ed.) (1972).
- Goodman, N. (1976), Languages of Art (Indianapolis: Hackett).
- Grice, H.P. (1957), 'Meaning', Philosophical Review, 66, 377-88.
- Grice, H.P. (1961), 'The Causal Theory of Perception', Proceedings of the Aristotelian Society, Supp. Vol. 1961, 121-52.
- Grice, H.P. (1968), 'Utterer's Meaning, Sentence-Meaning and Word-Meaning' in Foundations of Language, 4, 225-42.
- Gurney, E. (1971), 'Music as Impressive and Expressive' in Hospers (1971).
- Guttenplan, S. (ed.) (1975), Mind and Language: Wolfson College Lectures 1974 (Oxford: Clarendon Press).
- Hampshire, Stuart (1959), Thought and Action (London: Chatto and Windus).
- Hampshire, Stuart (1960), 'Feeling and Expression' in Hampshire (1972), Freedom of Mind (Oxford: Clarendon Press).
- Harrison, Bernard (1979), An Introduction to the Philosophy of Language (London: Macmillan).
- Hinde, R.A. (ed.) (1972), Non-Verbal Communication (Cambridge University Press).
- Hinde, R.A. (1975), 'The Comparative Study of Non- Verbal Communication' in Benthall and Polhemus (eds.) (1975).
- Hintikka, Jaako (1969), 'On the Logic of Perception' in Models for Modalities (Dordrecht: D.Reidel).
- Hooff, J.A.R.A.M. van (1972), 'A Comparative Approach to the Phylogeny of Laughter and Smiling' in Hinde (ed.) (1972).
- Hookway, C. and Pettit, P. (eds.) (1978), Action and Interpretation (Cambridge University Press).
- Hornsby, Jennifer (1980), Actions (London: Routledge and Kegan Paul).
- Hospers, John (ed.) (1969), Aesthetics (New York: The Free Press).
- Hospers, John (ed.) (1971), Artistic Expression (New York: The Free Press).
- Howell, Robert (1972), 'Seeing As', Synthese, 23, 400-22.
- Ishiguro, Hide (1967), 'Imagination', Proceedings of the Aristotelian Society, Supp. Vol. XLII.
- Jackson, Frank (1977), Perception (Cambridge University Press).
- Kenny, Anthony (1965), Action, Emotion and Will (London: Routledge and Kegan Paul).
- Kenny, Anthony (1976), Will, Freedom and Power (London: Routledge and Kegan Paul).

- La Barre, Weston (1947) 'The Cultural Basis of Emotions and Gestures', Journal of Personality, 16, 49-68. Reprinted in Polhemus (ed.) (1978).
- Lacan, Jacques (1949), 'Le stade du miroir comme formateur du fonction du je', Revue Francaise de Psychanalyse, 13.
- Leach, Edmund (1970), Levi-Strauss (London: Fontana).
- Leach, Edmund (1972), 'The Influence of Cultural Context in Non-Verbal Communication in Man', in Hinde (ed.) (1972).
- Levi-Strauss, Claude (1962), The Savage Mind. Translation (1966) (London: Weidenfield and Nicholson).
- Lewis, D.K. (1969), Convention, (Harvard University Press).
- Lewis, D.K. (1973), 'Causation', Journal of Philosophy, 70, 556-67. Reprinted in Sosa (ed.) (1975), Causation and Conditionals, (Oxford University Press).
- Lorenz, Konrad (1963), On Aggression. Translation by Marjorie Latzke (1966), (London: Methuen).
- McDowell, John (1978), 'On "The Reality of the Past"' in Hookway and Pettit (eds.) (1978).
- McDowell, John (1980), 'Meaning, Communication and Knowledge', in Z. van Straaten (ed.) (1980), Philosophical Subjects; Essays Presented to P.F.Strawson (Oxford: Clarendon Press).
- Merleau-Ponty, M. (1945 a), The Phenomenology of Perception. Translation by Colin Smith (1962), (London: Routledge and Kegan Paul).
- Merleau-Ponty, M. (1945 b), 'The Film and the New Psychology' in Merleau-Ponty (1948).
- Merleau-Ponty, M. (1947), 'The Primacy of Perception' in Merleau-Ponty (1948).
- Merleau-Ponty, M. (1948), Sense and Nonsense. Translation by Hubert L. and Patricia Allen Dreyfus (1964) (Northwestern University Press).
- Merleau-Ponty, M. (1960), 'The Child's Relations with Others'. Translation in Merleau-Ponty (1964 a).
- Merleau-Ponty, M. (1964 a), The Primacy of Perception and Other Essays (Northwestern University Press).
- Merleau-Ponty, M. (1964 b), 'Consciousness and the Acquisition of Language', Translation (1973) (Northwestern University Press).
- Peacocke, Christopher (1976), 'Truth Definitions and Actual Languages' in Evans and McDowell (1976).
- Peacocke, Christopher (1979), Holistic Explanation (Oxford:Clarendon Press).
- Pitcher, G. (ed.) (1966), Wittgenstein, The Philosophical Investigations: A Collection of Critical Essays (New York: Doubleday).

- Platts, M. (1979), Ways of Meaning (London: Routledge and Kegan Paul).
- Polhemus, Ted (ed.) (1978), Social Aspects of the Human Body, (Harmondsworth, Penguin).
- Quine, W. van Orman (1956), 'Quantifiers and Propositional Attitudes', Journal of Philosophy, 53, Reprinted in Quine (1976), The Ways of Paradox (Harvard University Press).
- Quine, W. van Orman (1960), Word and Object (Massachusetts Institute of Technology).
- Quine, W. van Orman (1969), 'Ontological Relativity' in Ontological Relativity and Other Essays (Columbia University Press).
- Ricoeur, Paul (1971), 'The Model of the Text: Meaningful Action Considered as a Text', Social Research, 38, 529-62.
- Ryle, G. (1949; 1970), The Concept of Mind (London: Hutchinson; Harmondsworth: Penguin).
- Sartre, J.P. (1939), Sketch for a Theory of the Emotions. Translation (1962), (London: Methuen).
- Sartre, J.P. (1945), Being and Nothingness. Translation by Hazel E. Barnes (1957), (London: Methuen).
- Scheler, Max (1913), The Nature of Sympathy. Translation by Peter Heath (1970), (Connecticut: Archon).
- Schiffer, S. (1972), Meaning (Oxford: Clarendon Press).
- Schwayder, D. (1965), The Stratification of Behaviour (London: Routledge and Kegan Paul).
- Searle, J.R. (1964), 'How to derive "ought" from "is"', Philosophical Review, 73, 43-58. Reprinted in Philippa Foot (ed.) (1967), Theories of Ethics (Oxford University Press).
- Searle, J.R. (1965), 'What is a Speech-Act?' in Max Black (ed.) (1965), Philosophy in America (London: Allen and Unwin).
- Searle, J.R. (1970), Speech Acts (Cambridge University Press).
- Sircello, G. (1972), Mind and Art (Princeton University Press).
- Stevenson, Charles L. (1944), Ethics and Language (Yale University Press).
- Stich, Stephen P. (1971), 'What Every Speaker Knows', Philosophical Review, LXXX, 476-96.
- Strawson, P.F. (1954), 'Wittgenstein's Philosophical Investigations', Mind, 63, 70-99. Reprinted in Strawson (1974), Freedom and Resentment (London: Methuen).
- Strawson, P.F. (1964), 'Intention and Convention in Speech-Acts', Philosophical Review, 73, 439-60. Page refs. to Strawson (1971).
- Strawson, P.F. (ed.) (1967), Philosophical Logic (Oxford University Press).

- Strawson, P.F. (1970), 'Meaning and Truth' in Strawson (1971).
- Strawson, P.F. (1971), Logico-Linguistic Papers (London: Methuen).
- Taylor, Charles (1979), 'Action as Expression' in Cora Diamond and Jenny Teichman (eds.) (1979), Intention and Intentionality: Essays in Honour of G.E.M. Anscombe (Sussex: Harvester).
- Tormey, Alan (1971), The Concept of Expression (Princeton University Press).
- Unger, Peter (1972), 'Propositional Verbs and Knowledge', Journal of Philosophy, LXIX.
- Urmson, J.O. (1952), 'Parenthetical Verbs', Mind, LXI, 480-96.
- Urmson, J.O. (1968), The Emotive Theory of Ethics, (London: Hutchinson).
- Vesey, G.N.A. (1955), 'Seeing and Seeing As', Proceedings of the Aristotelian Society, 56, 109-24.
- Warnock, Mary (1976), Imagination (London: Faber).
- Wittgenstein, Ludwig (1958), Philosophical Investigations (Oxford: Blackwell).
- Wollheim, Richard (1964), 'On Expression and Expressionism', Revue Internationale de Philosophie, nos. 68-9, fascs. 2-3, 270-89.
- Wollheim, Richard (1966), 'Expression' in Royal Institute of Philosophy Lectures 1966-67, Vol. I : The Human Agent, ed. G.N.A. Vesey (1968) (London). Page refs. to Wollheim (1973).
- Wollheim, Richard (1968; 1970), Art and its Objects (New York: Harper and Row; Harmondsworth: Penguin).
- Wollheim, Richard (1970), 'Nelson Goodman's Languages of Art', Journal of Philosophy, LXXII, 531-9. Page refs. to Wollheim (1973).
- Wollheim, Richard (1973), On Art and the Mind (London: Allen Lane).
- Wollheim, Richard (1980), 'Seeing-as, Seeing-in and Pictorial Representation' in Art and its Objects, 2nd Edition (Cambridge University Press).
- Ziff, Paul (1958), 'About Behaviourism', Analysis, XVIII. Page refs. to Ziff (1966), Philosophic Turnings (Cornell University Press).