

AN ANTICIPATION OF REVERSAL THEORY FROM WITHIN A CONCEPTUAL-  
ANALYTIC AND BEHAVIORIST PERSPECTIVE.

U.T. Place

School of Psychology  
University of Wales at Bangor  
UK

School of Philosophy  
University of Leeds  
UK

J.A. Wheeler Vega\*

Department of Psychology and Research in Education  
University of Kansas  
Bailey Hall  
Lawrence  
KS 66045-2330  
USA

*Paper presented at the Ninth International Conference on Reversal Theory,  
June 28 – July 2, 1999, University of Windsor, Ontario, CA.*

\*Contact second author with all inquiries about this paper.  
EMAIL [wheeJ775@newschool.edu](mailto:wheeJ775@newschool.edu)

## Abstract

Michael Apter denies that behaviorism can provide an adequate account of human action, referring to it in one place as “a kind of methodological vandalism” (Apter, 1989, p.2). It is the purpose of this paper to show how the first author came, as a behaviorist and analytic philosopher, to a position that anticipates reversal theory to a surprising extent.

The basis of this position is an analysis of polar statements concerning ‘wanting’: ‘X wants O’, and ‘X does not want O’. These sentences imply a number of corollaries. For example, if ‘X wants O’, then: ‘X will be *pleased* if O appears’, ‘X will be *worried* if it looks as if O will not appear, and ‘X will be *angry* or *miserable* if O fails to appear’. Contrasting entailments follow ‘X does not want O’. These implications display the relationship between the motivational concepts of ‘wanting’ and ‘not wanting’, and emotion concepts such as being pleased, worried, angry, miserable, &c. This set of reciprocally related entailments provide, it will be argued, the conceptual foundation of reversal theory.

This analysis led the first author to develop a behavioral theory of emotion, in which the various emotions can be located on two dimensions (after Myers, 1923): ‘pleasant/unpleasant’, and ‘high-arousal/low-arousal’. Emotions are distinguished by reference to a third variable: a characteristic ‘impulse’ appropriate to the type of contingency in which the emotion in question is evoked. The notions of ‘wanting’ and ‘not wanting’ are defined, in the language of operant psychology, as differences in the reinforcing effect of actual and potential stimuli with respect to actual and potential operant responses by the organism.

Some illustrative clinical and experimental applications of the theory by the first author, in the 1960’s, will be outlined.

## INTRODUCTION

In his first book on reversal theory (1982) Professor Apter located the theory as an example of a wider approach that he calls *structural phenomenology*. He suggests: "It would be quite possible to disagree with the assertions of the theory while supporting the general approach." (p. 331).

My taste for irony is excited by another possibility: one might agree with the basic assertions of the theory while rejecting the general approach. I hope to provide an exposition of U. T. Place's theory of emotions, developed in the mid 1970's, which shows that his performance centered theory is quite compatible with most of reversal theory, and anticipates some of its central features. I will also argue that an adequate theory of motivation and emotion cannot be principally based on subjective experience.

## ANALYSING EMOTION

Place distinguishes three questions that can be asked about emotions: (1) What do emotion words mean? (2) What is it that produces a particular emotional reaction in a person? (3) What does an emotion consist in when analyzed in micro-reductive terms? (1974a, p.1)

He suggests that the first question: "What are we saying about someone when we describe him as relieved, pleased, excited, angry, frightened, distressed, miserable or bored?" (1974a, p.3) must be addressed before questions about the causes of particular emotions, or about their physiological basis. Instead of taking for granted our everyday conceptions of emotion and their attribution and using them in psychological theory without question, some analysis of these terms is required.

One everyday assumption about emotions is common to other subjective experiences: their reporting is taken to be incorrigible; one's first-person authority is usually sacrosanct. But this assumption ignores everyday practice with regard to the attribution of emotions:

If a man exhibits the kind of behaviour that is characteristic of someone who is pleased, angry or afraid, we are perfectly justified in concluding that he *is* pleased, angry or afraid regardless of what he may say about the experiences he is currently having. (ibid., p.11)

This conceptual primacy of behaviour over experience in the definition of emotional states is also demonstrated by the fact that when we try to describe the experiences that are characteristic of a given emotional state, we find ourselves having to distinguish between those glows, thrills, surges and thuds which are characteristic of emotional states

and other glows, thrills, surges and thuds which are mere sensations without emotional significance; and in order to do this we have to specify the emotional state of which the experience in question is characteristic, which in turn implies that the emotion is something (behavioral) over and above the experience that is typical of it. (ibid., p.12)

This is a development by Place of a nice remark by Gilbert Ryle (1949/1976) in *The Concept of Mind*:

It is an important linguistic fact that these names for specific feelings, such as 'itch', 'qualm', and 'pang' are also used as names of specific bodily sensations. If someone says that he has just felt a twinge, it is proper to ask whether it was a twinge of remorse or of rheumatism. (p. 82)

These points imply that a theory of emotion centered on subjective experience is incomplete. Emotional experiences are therefore described by Place as, "characteristic, rather than constitutive of emotional states." (1974a, p.12)

## EMOTION AND EXPERIENCE

It is now common in psychology to identify three components of emotion: the subjective or experiential, the behavioral, and the physiological. (Gross, 1987; Place, 1974a). While RT is centrally concerned with the experience of emotion, and with the hypothesized deep structures that determine these surface subjective phenomena, Place emphasizes the centrality of behaviour.

Professor Apter does of course recognize that behaviour is an important element in psychological theory. As he says:

Since part of the phenomenal field will be the individual's own actions, structural phenomenology is concerned with behaviour as well as mentation...however, the starting point and the pivotal centre of interest is always the phenomenal field. (1989, p.3)

But he is too quick to discard 'behaviorism'. He says that he rejects, "any approach...which sees experience as no more than a by-product or side issue...[and is] thus totally opposed to behaviorism in any of its forms, regarding it as a kind of methodological vandalism." (ibid., p. 2). But behaviorism comes in significantly different forms, and his repeated dismissal of a diverse range of theories and approaches is unsatisfactory.

This leads me to an important methodological aside, as I think that this an ironic complaint of behaviorism. In his first book on reversal theory

(1982), he claims that the metamotivational states which appear in his theory (excitement, boredom, etc.) are *mental epiphenomena* of physical but unspecified metamotivational systems, such as the telic and paratelic. The existence of these systems is inferred from introspected conscious experience. And in his 1989 book he returns even more clearly to the confident introspectionism of Descartes. He says:

The development of...[reversal] theory itself is 'experience driven'...[T]he theory derives not from previous research results, or the development of some new type of instrumentation, or certain abstract ideas...Rather, the principal starting point is that cool, clear source of everything we know about ourselves: our own consciousness. The general approach of structural phenomenology is therefore not just 'top-down' but is also what one might call 'inside out'. (p. 7)

Two key assumptions of structural phenomenology are revealed in these passages:

#### OVERHEAD

- (1) Epiphenomenalism is true.
- (2) It is valid to infer the deep structure of motivation from the surface phenomena of emotional experience.

Both of these moves cause problems:

- (1) A considerable literature has developed about 'epiphenomenalism' since the idea was articulated by T.H. Huxley in 1874 (1899), and a number of stronger and weaker versions have been distinguished (McLaughlin, 1994). The general idea is of an effect without causal powers. Epiphenomenalism about mental states is far from an obvious assumption. It is, on the contrary, poison to most kinds of psychology.

Place has recently argued (1999) that if epiphenomenalism is true, then first-hand introspective reports (such as those on which RT focuses) are no such thing, because a first hand report, to be first hand, must be causally related to the event which it reports. But, if mental events don't cause anything, then they can't cause our reports of subjective experiences. Epiphenomenalism therefore entails, rather ironically for Apter, a radical behaviorism, as our subjective reports would then come out as some kind of confabulation, which we would need not heed.

The endorsement of epiphenomenalism also disqualifies RT from being a kind of *action theory*, as Apter has suggested it is, where action is understood as (in his words), "behaviour plus subjective

meaning...especially the intentions of the person performing the behaviour." (1989, p.3). It follows that if our conscious intentions are epiphenomenal, they cannot contribute causally to action, and people are therefore not the free agents that action theories depict.

- (2) As William James noted, introspection provides us, at best, only with the 'perchings' and not the 'flights' of our mentation (Valentine, 1978), the products not the processes. To infer process from product is daring and difficult. Somewhat like Freud's later theorizing about the mind, Professor Apter's inference of metamotivational processes from the introspection of subjective experiences is clever and bold but insecurely grounded.

## THEORIES OF EMOTION

With these observations in mind, I now turn to the development of Place's theory of emotion. Place notes that Wundt's (1896) early analysis of emotion made use of three dimensions: *pleasantness-unpleasantness*, *excitement-depression*, and *strain-relaxation*. He notes that excitement-depression and strain-relaxation are not orthogonal dimensions, as excitement and depression have connotations of pleasantness and unpleasantness respectively.

Place therefore follows the British psychologist C. S. Myers in his theory, who in 1923 developed a two-dimensional analysis of emotions. Myers held that there were two basic states of an organism, strain (S) and relaxation (R) that interact with two basic environmental states, either favorable (F) or unfavorable (UF). These in combination yield four primitive affects, or emotional states.

### OVERHEAD

SF – *exhilaration* [EXCITEMENT], *gladness, and interest*

SUF – *uneasiness* [ANXIETY], *distress, and repugnance*

RF – *ease* [RELAXATION], *bliss, contentment*

RUF – *depression, sadness, and apathy* [BOREDOM]

(Myers uses slightly different abbreviations, but I prefer the accidental humour of mine.)

It can be seen how closely these parallel, at least *prima facie*, some of the basic emotion categories of reversal theory: excitement, anxiety, relaxation, and boredom. There is also a suggestion by Myers that some emotions straddle these categories.

Place develops Myers' account into a picture of emotions as ranged on two dimensions: pleasant-unpleasant and high arousal-low arousal. His

account can also be contrasted with the categories of Professor Apter's theory.

OVERHEAD

	PLEASANT HEDONIC TONE	MIXED HEDONIC TONE	UNPLEASANT HEDONIC TONE
HIGH AROUSAL	<i>Excitement</i> Sexual arousal	Anger Awe	Fear <i>Anxiety</i> Distress
MEDIUM AROUSAL	Pleasure Enjoyment Love	Nostalgia Pity	Disgust Shame Embarrassment
LOW AROUSAL	Relief <i>Relaxation</i>	Weariness <i>Apathy</i>	Grief Misery Depression

*Note: I think there is some question about the arousal level of unpleasant states such as grief and depression, which may actually involve the organism in states of high physiological arousal.*

The standout differences between Apter's and Place's categories are *depression, apathy, and boredom*. However, the fine differences between the two theories are not really the main issue of my paper. There are deep differences and broad agreements that deserve more attention. For example, arousal is defined subjectively by Apter (as 'felt arousal'), but objectively by Place by reference to a range of physiological measures.

MEANING AND IMPULSES

Another profound issue concerns the meaning of emotions. Professor Apter emphasizes the focus of RT on the meaning of individuals' experience. He says that RT, "starts from subjective experience and interprets behaviour, or physiological processes, in the light of this experience." (1989 p.7), and his notion of 'felt significance' follows from this emphasis.

Contrastingly, Myers had a notion of the development of feelings that encourages a behaviorist account of the meaning of emotions. Myers suggested that our feelings were selected by virtue of their direct relationship to the complex reactions of the organism to demanding stimuli:

[W]ith the progress of mental and bodily evolution a generalised form or pattern of visceral and somatic reaction...becomes associated with a definite kind of...affect. The affect comes largely to be the meaning of the reaction to the self. (1923, p. 7)

The special emotion feeling is dependent...on the general character of the situation which confronts the subject, and on the common purpose of the alternative *instinctive reactions*. Fear and disgust, for instance, are typically connected with escape and rejection, anger, lust with hostile attack and sexual possession, respectively. (my italics, *ibid.*, p. 9)

Thus, feelings become the meaning of crucial behaviour. It can also be seen that particular emotions have both a characteristic range of feelings and a characteristic range of behavioral reactions.

This notion is considerably developed by Place. Although the two dimensions of hedonic tone and arousal order emotions quite well, some more specificity is required. As argued above, subjective experiences are not sufficient to distinguish emotions, so Place proposes a third variable, although not a dimension as such. This differentiating variable, between, say, disgust and embarrassment is the performative impulse that attends each emotion.

Disgust and embarrassment are both highly unpleasant, and involve moderate arousal. But the characteristic behavioral impulse for disgust is vomiting, or less dramatically, avoidance of the noxious stimulus. Embarrassment characteristically involves hiding, or, less extremely, various forms of social withdrawal.

Myers speaks of 'instinctive reactions' and this is partly right, but needs some elaboration. Place defines the role of an impulse as follows:

The effect of an emotional response is to establish a condition under which *behaviour of a particular kind* becomes intrinsically reinforcing. In the case of anger any response which causes pain or damage to the object of the anger (which of course may not be its cause) will be powerfully reinforced. (my emphasis; U. T. Place, personal communication, May 15, 1999)

Impulses are precisely adapted to the kind of contingency which elicits them, so that positive contingencies elicit impulses which serve to secure or maintain those conditions, whereas negative contingencies evoke impulses whose function is the avoidance or cessation of the condition.

Place asserts that a great many human emotions are common to lower animals, and so, as Myers also suggests, the link between emotional responses and *particular kinds of behaviour*, a link which serves to enhance the sensitivity to reinforcement of relevant behaviors, was forged in our evolutionary past.

Place also makes the important point (motivated by Wittgenstein's discussion of the learning of sensation words (1968, §244)) that: "Verbal



impulses (exclamations) tend to replace motor impulses as expressions of emotion in adult humans." (1974a, p.20). So, we should be most attentive to verbal emotional behaviour, and also not expect to be able to point out any specific behaviour as indicative of some emotion.

Both Skinner and Ryle also recognized this latter point. As Ryle says, with the examples of being sulky or hilarious:

[N]o one of...[the characteristic responses] is a necessary or sufficient condition of being sulky or hilarious. (1949/1976, p.93)

Skinner too notes a range of typical responses for each emotion:

The "angry" man shows an increased probability of striking, insulting, or otherwise inflicting injury...The man "in love" shows an increased tendency to aid, favor, be with, and caress. (1953, p.162)

It is therefore a characteristic kind of impulse, such as 'to injure', and not any particular behaviour that is specific to each emotion.

This seems better, but still not quite right. In Skinner's example, *love* can be seen to have a range of impulses: 'to aid, favor, be with, caress'. It seems then, that some emotions may have several characteristic impulses. But only a few, and always some and not others. We would never say that the impulse to caress was characteristic of *anger*, or that assault was a characteristic impulse of *love*. Our natural language categories for emotions have clearly resolved into fairly manageable units, with only one or a few typical impulses for each emotion category.

To return now to the meaning of emotion, rather than working 'outwards' from 'inner' subjective feelings to the interpretation of behaviour, as Apter suggests, we can say that the meaning of emotions is built in. If, as Myers suggests, and I think my discussion supports, feelings are the meaning of crucial behaviour, this also implies an explanation of the link between emotion and motivation. For what is common to almost all impulses is their achievement of some adaptive goal for the organism, in virtue of the links established between impulse and crucial behaviour by our ancestors.

## EMOTION AND MOTIVATION

A theory of emotion should give some account of the relationship between emotion and motivation. A central feature of RT is Apter's argument against the limitations of homeostatic drive theories of motivation. Instead he posits multi- or bi-stable motivational systems that actively seek high or low arousal at different times, rather than some optimal median level.

But the link between emotion and motivation could still be clearer, and the previous discussion suggests one way to make it so. As Place explains:

The biological functions of emotional reactions are closely related to the biological function of the motivational process, which has the function of securing the sequential organization of behaviour in such a way that specific positive contingencies likely to increase the survival chances of the individual and the species are brought about, while negative contingencies constituting a threat to survival are avoided. The biological emergencies for which it is a function of the emotion to mobilise the appropriate behaviour are always situations towards or away from which it is the function of the motivational process to organise behaviour. (1974b, p.5)

This leads to the observation that:

In ordinary language there is a conceptual relationship between the motivational concepts of 'wanting' and 'not wanting' and...emotion concepts...[such as] 'being pleased, distressed, worried, frightened, angry, miserable and relieved'. (op. cit.)

#### OVERHEAD

It would seem to be the case that to say that someone *wants* something entails that he will be:

- (a) *pleased*, if he thinks he has got or is about to get what he wants,
- (b) *worried* or *afraid*, if he thinks that he may not get it,
- (c) either *angry* or *depressed*, if he thinks he no longer has any hope of getting it.

Similarly, to say someone *does not want* something entails that he will:

- (d) be *worried* or *afraid*, if he thinks it is likely to come about,
- (e) *angry* or *distressed*, if it does come about or if he thinks it has,
- (f) and *relieved*, if he thinks it has not come about or is not now likely to come about.

As suggested by the previous discussion, this analysis of the link between motivation and emotion can also be seen to extend from motivation to behaviour:

[I]t also seems correct to say that to say of someone that he is *pleased by* or *relieved at* something entails that he does not or would not want it to be otherwise, that to say that he is *afraid of* something

entails wanting to escape from, avoid or otherwise prevent it from happening, to say that he is *angry with* someone or something entails wanting to attack, hurt or destroy that something, that to say that he is *distressed at* something is to say that he wants to escape from the circumstances which are the occasion of his distress, while to say that he is *depressed* or *miserable* is to say that there is little or nothing that he positively wants to do and much that he positively does not want to do. (1974a, p.16)

## REVERSAL

A key element in RT, so far absent from my discussion of Place's theory, is the notion of reversal. It is to this subject that I finally turn, with a discussion of some experimental work with clinical subjects conducted by Place in the mid-1960's.

Place's experiments sought to test two competing accounts of the effect of differences in mood on motivation. It will be no surprise that an operational definition of motivation was required for such experiments: operant responding was the obvious choice.

One account was derived from remarks by Skinner. Skinner suggested a simple direct relationship between mood and operant responding, such that:

[E]lation consists in a general increase in the rate and probability of operant response emission and...depression consists in a corresponding general reduction in the rate and probability of operant response emission. (Place, 1968, p.2)

The second account was derived from remarks by Ryle:

If a person is too gay to brood over a rebuff, he is not undergoing so violent a feeling that he can think of nothing else, and therefore not of the rebuff; on the contrary, he enjoys much more than usual all the things he does and all the thoughts he thinks, including thoughts of the rebuff. He does not mind thinking of it as much as he would usually do. (1949/1976, p.97)

As [Place] interpret[s] it, what Ryle is saying here is that a mood of elation or happiness consists of two things,

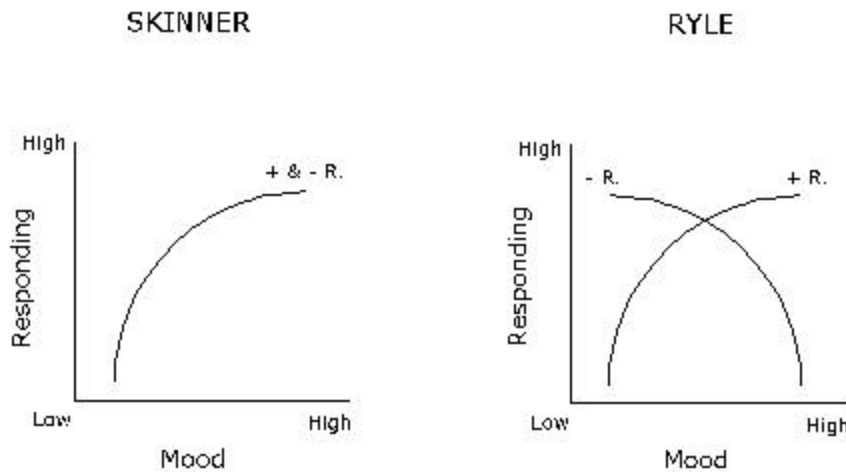
- (a) an increase in the individual's capacity of enjoyment [sensitivity to positive reinforcement], and
- (b) an overall reduction in the individual's sensitivity to distress [sensitivity to negative reinforcement].

Ryle makes no specific mention of depressed or unhappy frames of mind in this connection; but since depression is presumably the opposite of elation, it would seem to be a corollary of his view that depression consists in,

- (c) a reduction in the individual's capacity for enjoyment [sensitivity to positive reinforcement], and
- (d) an increase in the individual's sensitivity to distress [sensitivity to negative reinforcement] (Place, 1968, p.3)

The relationship between mood and motivation (operant responding) implied by Ryle's comments is therefore more complex than on Skinner's account.

### OVERHEAD



The rate of responding under conditions of positive reinforcement would be expected to follow the same course as predicted by Skinner...rising above normal in elation and falling below normal in depression. The rate of responding in conditions of negative reinforcement, on the other hand, would be expected to do the opposite, to fall below normal in elation and rise above normal in depression. (ibid., p.5)

In RT, reversals are said to occur between metamotivational systems, for example, between the arousal avoiding or telic system, and the arousal seeking or paratelic system. For Place, the notion of reversal refers to the change in sensitivity to reinforcement effected by changes of mood, such that high and low moods have opposite effects under different environmental contingencies.

## Method

To test these competing hypotheses, Place refined a previously used instrument comprising a Morse code key, reinforcement counter, and red and green lamps to indicate negative and positive reinforcement respectively.

When the green light is on, the reinforcement is positive. In this condition pressing the key will cause the counter to start counting up. When the red light is on, the reinforcement is negative and the counter counts down unless the subject responds so as to stop it. The movement of the counter derives its value as a positive reinforcer and aversive stimulus respectively from a cash payment that is made at the end of the experimental session. (ibid., p.8)

## Subjects

Subjects were two, a man and a woman, both psychiatric patients diagnosed with manic depression, and both with histories of unusually regular and predictable fluctuations of mood between elation and depression.

## Results

For the sake of brevity, I must pass over some of the interesting interpretation of the data and just present the theoretically relevant conclusions. In short, most of the predictions of the Rylean account were supported. These were:

1. The hypothesis common to the Skinnerian and Rylean accounts that responding would be positively related to mood under positive reinforcement.
2. The Rylean hypothesis that under negative reinforcement the relationship between mood and responding would be reversed, so that responding would be lower in an elated, and higher in a depressed, mood.

These results do not imply the operation of different motivational systems that alter mood when active, but distinctly different motivational reactions to contrasting environments within the same mood range.

## References

- Apter, M. J. (1982). The Experience of Motivation: The Theory of Psychological Reversals. London: Academic Press.
- Apter, M. J. (1989). Reversal Theory: Motivation, Emotion, and Personality. London: Routledge.
- Gross, R. D. (1987). Psychology: The Science of Mind and Behaviour. London: Hodder and Stoughton.
- Huxley, T. H. (1899). On the hypothesis that animals are automata, and its history. In, Method and Results: Essays (pp. 199-250). New York: D. Appleton & Co. (Paper dated 1874).
- McLaughlin, B. P. (1994). Epiphenomenalism. In S. Guttenplan (Ed.), A Companion to the Philosophy of Mind (pp. 277-88). Oxford: Blackwell.
- Myers, C. S. (1923, March). The evolution of feeling. Australian Journal of Psychology & Philosophy, 1(1), 3-11.
- Place, U. T. (1968). The use of operant responding as a measure of mood fluctuation in periodic psychosis. Paper presented at the Annual Conference of the Experimental Analysis of Behaviour Group, Nottingham, UK.
- Place, U. T. (1974a). Emotion concepts and learning theory 1: Theories of emotion and the nature of emotional reactions. Lectures on the Metaphysical Foundation of Empirical Psychology, Lecture 25. Amsterdam, 1973-4.
- Place, U. T. (1974b). Emotion concepts and learning theory 2: The biological functions of emotional reactions and their elicitation conditions. Lectures on the Metaphysical Foundation of Empirical Psychology, Lecture 26. Amsterdam, 1973-4.
- Place, U.T. (1999, May) Are qualia dispositional properties? Paper presented at the Università di Siena and Terza Università di Roma conference on "Consciousness Naturalized", Certosa di Pontignano, Siena.
- Ryle, G. (1976). The Concept of Mind. Harmondsworth: Penguin. (First published 1949).
- Skinner, B. F. (1953). Science and Human Behavior. New York: Free Press.
- Valentine, E. R. (1978). Perchings and flights: introspection, In J. Radford & A. Burton (Eds.), Thinking in Perspective, (pp. 1-22). London: Methuen.
- Wittgenstein, L. (1968). Philosophical Investigations (3rd ed). (G. E. M. Anscombe, Trans.). Oxford: Blackwell.
- Wundt, W. (1896). Grundriss der Psychologie [Outlines of Psychology]. Leipzig: Engelman.