British Journal of Psychology (1989), **80**, 509–515 *Printed in Great Britain*

© 1989 The British Psychological Society

Human reactions to reward and punishment: A questionnaire examination of Gray's personality theory

Glenn D. Wilson*, Paul T. Barrett and **Jeffrey A. Gray**

Institute of Psychiatry, University of London, De Crespigny Park, London SE5 8AF, UK

The development of the Gray–Wilson Personality Questionnaire is described; this is an instrument designed to measure human equivalents of six animal behaviour paradigms – Approach, Active Avoidance, Passive Avoidance, Extinction, Fight and Flight. Although these six scales showed satisfactory internal consistency they failed to link up into the three major systems suggested by Gray's personality theory. The strongest associations were between Fight and Approach and between Flight and Passive Avoidance. This raises questions as to how the neurological systems of activation, inhibition and fight/flight are related to human personality structure.

Gray (1987*a*, *b*) has proposed a three-factor model of personality based on animal learning paradigms, drug effects and neuropsychological studies. The three main personality dimensions relate to basic brain/behavioural mechanisms as follows.

(1) The *activation system* deals with behaviour elicited by rewards, or more precisely, incentives, the conditioned signals of reward. This has two aspects: (*a*) responsiveness to stimuli in the environment that are associated with primary rewards (eating, drinking, copulation, etc.), which we shall call *approach*, and (*b*) instrumental behaviour which reduces the likelihood of punishment when there are signals in the environment that punishment might occur, known in laboratory studies as *active avoidance*.

(2) The *inhibition system* organizes responses to conditioned signals of punishment (and frustrative non-reward). Its main effects are inhibition of ongoing behaviour, increased attention and increased arousal. Because these reactions are diminished by anxiolytic drugs, it may be presumed that its operation is accompanied by the emotional state called 'anxiety'. Two laboratory phenomena thought to be manifestations of the inhibition system are (*a*) *passive avoidance* (reducing risk of punishment by inactivity and submission) and (*b*) *extinction* (readily abandoning behaviours that are not rewarded).

(3) The *fight/flight system* mediates the behavioural effects of unconditioned aversive events (again including punishment *per se* and frustrative non-reward). Such behaviours include (*a*) *flight* (i.e. rapid escape from the source of punishment) and (*b*) *fight* ('defensive' aggression, as distinguished from predatory aggression).

These three behavioural systems are conceived as virtually independent from one

* Requests for reprints.

another, with separate neurological mechanisms. The prototypes are also held to have clinical relevance; for example, phobias are seen as a form of passive avoidance, and depression may be related to extinction (if one accepts Lewinsohn's, 1975, view of depression as deriving from insufficient reinforcement).

This three-factor theory is an expansion of an earlier two-factor model of personality which rotated Eysenck's extraversion–introversion and neuroticism factors into a near diagonal position (corresponding with anxiety and impulsiveness) and identified them with sensitivity to punishment and reward respectively (Gray, 1981). A major way in which this theory differed from that of Eysenck is that, whereas Eysenck proposes that introverts (with their chronically higher levels of cortical arousal) are generally more conditionable than extraverts, Gray's theory predicts that extraverts will condition more easily than introverts under conditions of reward (as opposed to the more common paradigm in which the unconditioned stimulus is a punishing event such as electric shock). There are now several studies which support this prediction (Boddy, Carver & Rowley, 1986; Nichols & Newman, 1986; Patterson, Kosson & Newman, 1987; Torrubia & Tobena, 1984). Although the three-factor version of Gray's theory integrates a great deal of research on animal learning and the psychophysiology of emotions and, intriguingly, raises the number of major personality dimensions to three as Eysenck has also done, it has not yet been tested at the human level.

Given the potential importance of this theory in linking animal laboratory paradigms to clinical phenomena, it was thought worthwhile to attempt to develop a human personality questionnaire derived explicity from it. The main hypothesis was that scales designed to measure human equivalents of the six laboratory phenomena listed above would link up into pairs corresponding with Gray's three major proposed neuropsychological systems.

Method

Ten items were devised to tap each of the six laboratory-established types of behaviour (approach, active avoidance, passive avoidance, extinction, fight and flight). An attempt was made to vary the particular content of the items within each scale. For example, approach was variously represented as towards money, food, addictive substances, general consumer commodities, sex partners, pleasurable social events, attractive clothing, birthday presents, theatrical experiences and career advancement. To reduce the likelihood that agreement response bias would influence scale scores each of the 10 items was then matched with an approximate logical reversal of itself. This yielded 20 items in each scale, half scored positively and half negatively. The full Gray–Wilson Personality Questionnaire with instructions and scoring details is available from the authors on request.

Subjects were recruited in two ways. First, the parents of children at a South-East London primary school were approached while waiting to collect their children and asked if they would mind completing the questionnaire at home and returning it by mail to the researchers. Second, students in halls of residence at a South London Technical and Liberal Arts College were left questionnaires in their mail boxes, with a similar request for help and a reply-paid envelope. All subjects were offered anonymity if they wished, but they were also asked if they would mind giving a contact address for possible follow-up studies. In any case, demographic information such as age, sex and occupation was collected. Subjects were also asked to complete the Eysenck Personality Questionnaire (Eysenck & Eysenck, 1975).

In all, 750 questionnaires were distributed and 243 were returned (159 women and 84 men). This

return rate is fairly typical for a study of this kind. The mean age for women was 25.46 years (SD = 8.78) and for men 26.71 (SD = 8.16).

It is not claimed that this sample is perfectly representative of British adults at large, but it is broad in scope and contains sufficient variability for the correlational and factor analysis to be described.

Results

Means, SDs and alpha coefficients for the six Gray–Wilson scales are shown in Table 1. These scores are conveniently arranged around the mid-point of 20, thus maximizing scale discriminability. Alpha coefficients suggest reasonable levels of internal consistency, the exception being the Active Avoidance scale as applied to female subjects, where alpha falls as low as 0.35.

While it might be argued that these alpha coefficients are inflated by the inclusion of pairs of items that are approximate logical reversals, this criticism would carry little weight. All scales with high internal consistency contain items of overlapping content, and it is worse if they are all scored in the same direction. Many scales artificially enhance their apparent factorial purity by confounding response sets with item content (Wilson, 1975). In any case, the alpha coefficient is best regarded as a measure of item redundancy: alphas that are too high imply narrowness and tautology. Comparisons between men and women, using *t* tests, show that women are significantly higher on the Active Avoidance and Flight scales. There were no

Table 1. Means, SDs and alpha coefficients for six GWPQ scales

	Men (n = 84)			Women (n = 159)			
	Mean	SD	alpha	Mean	SD	alpha	*P* diff
Approach	17.46	6.91	0.71	17.58	6.66	0.68	n.s.
Active Avoidance	22.71	6.05	0.61	25.69	4.53	0.35	< 0.001
Passive Avoidance	17.40	5.96	0.58	18.3	5.89	0.59	n.s.
Extinction	18.81	5.97	0.61	20.31	6.10	0.63	n.s.
Fight	18.19	6.43	0.65	18.27	6.97	0.71	n.s.
Flight	16.94	6.33	0.65	19.85	6.59	0.71	< 0.01

significant sex differences on the other four scales.

Table 2 shows intercorrelations among the six scales for the female sample, and for males in parentheses. Generally, there is a high degree of similarity between men and women as regards these correlations. Fight and Approach are substantially related, as are Flight and Passive Avoidance. Approach is negatively related to Active Avoidance and there are significant correlations between Extinction and Passive Avoidance, Approach and Passive Avoidance, and (negatively) between Fight and Active Avoidance.

These relationships are further illustrated in a similarity space analysis (Figure 1), which also includes EPQ variables. This analysis, known as the Guttman–Lingoes

Table 2. Intercorrelations among the six GWPQ scales for female subjects (males in parentheses)

	Approach	Actavoid	Passavoid	Extinct	Fight
Actavoid	−0.25 (−0.38)				
Passavoid	0.23 (0.20)	0.00 (−0.01)			
Extinction	0.16 (−0.04)	0.01 (0.17)	0.26 (0.37)		
Fight	0.44 (0.45)	−0.28 (−0.25)	0.11 (0.16)	0.11 (−0.23)	
Flight	0.10 (0.19)	0.10 (0.14)	0.48 (0.36)	0.14 (0.10)	0.13 (0.16)

Key. Actavoid = Active Avoidance; Passavoid = Passive Avoidance.

Table 3. Correlations between EPQ and GWPQ scales for female subjects (males in parentheses)

	EPQ scales			
GWPQ scales	E	N	P	L
Approach	0.27 (0.22)	0.12 (0.14)	0.39 (0.32)	−0.33 (−0.28)
Actavoid	−0.03 (−0.23)	0.10 (0.08)	−0.42 (−0.36)	0.17 (0.17)
Passavoid	−0.21 (−0.19)	0.42 (0.67)	0.03 (0.18)	−0.15 (−0.26)
Extinction	−0.35 (−0.42)	0.20 (0.26)	0.10 (0.04)	−0.00 (0.12)
Fight	0.22 (0.21)	0.14 (0.34)	0.40 (0.30)	−0.42 (−0.31)
Flight	−0.09 (0.09)	0.18 (0.37)	−0.23 (0.04)	−0.10 (−0.34)

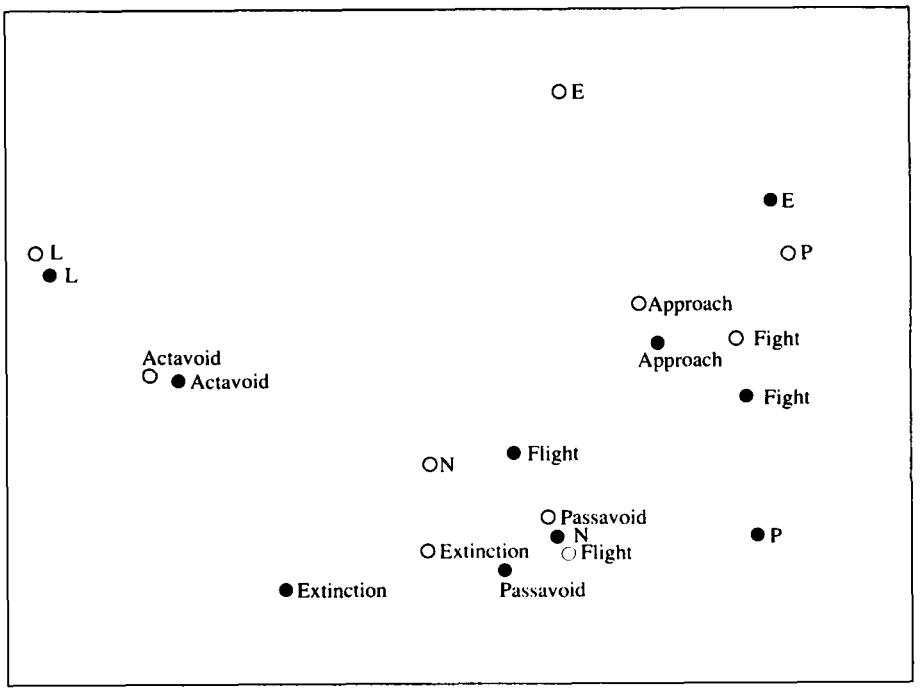

Figure 1. Spatial representation of the similarity among six Gray–Wilson and four Eysenck personality scales (E, L, N, P). ●, males; O, females.

Smallest Space Analysis (SSA), is a non-linear multidimensional scaling technique (Lingoes, 1973). Goodness-of-fit measures suggested that for both males and females a two-dimensional representation was permissible ('stress' indices of 0·17 and 0·19 respectively). Again Fig. 1 reveals a remarkable similarity between the results for men and women.

Table 3 shows correlations between Gray–Wilson scales and EPQ scores. Approach is most strongly related to P, but also goes with E and low L scores. Active Avoidance goes with low P, and introversion for men only. Passive Avoidance is strongly related to N, with correlations with introversion and low L of borderline significance. Extinction goes with introversion and high N. Fight is related to high P, low L and a tendency to extraversion, and Flight goes with N (particularly in men) and low P (only in women). These correlations are consistent with the configurations shown in Fig. 1.

Discussion

These results present a difficulty for Gray's theory of personality as applied to human subjects. Scales were explicitly designed to measure human equivalents of six animal behaviour paradigms and these proved acceptable as regards normal psychometric criteria. However, the intercorrelations among them were not in accord with Gray's theory.

The most contradictory result was the finding of a significant negative relationship between Approach and Active Avoidance, when according to Gray's theory these are both aspects of Activation (and therefore ought to be positively correlated). Also, within the framework of Gray's theory, Fight and Flight are manifestations of the same neuropsychological system, but the correlation, while positive, failed to reach significance. Only the predicted relationship between Passive Avoidance and Extinction (as two major parts of the Inhibition system) was confirmed in this study.

Intuitively, these correlations do not seem surprising. The kind of person who moves strongly towards rewards such as money, food and sex (the Approacher) is not usually careful to avoid punishment by lawful parking, medical check-ups and carrying an umbrella (the Active Avoider). One trait sounds like impulsiveness and the other compulsiveness, and these two do not go together clinically. Why behaviours of these two types should depend on a common neural system in rats yet be negatively correlated in humans is something of a puzzle.

Similarly, although Fight and Flight are no doubt managed by the same neurohormonal 'emergency' system, what is more interesting at the human level is why one person characteristically chooses to fight when subjected to threat and another to flee. One person we call 'aggressive' and the other 'fearful' and, although both may be 'emotional', we think of them as rather different types of people. Thus Gray's theory that the Fight/Flight system underlies an important dimension of temperament requires further examination at the human level of analysis.

The sex differences in Table 1 should also be noted in this connection. While there is no difference between men and women on Approach, women are significantly higher on Active Avoidance (confirming the impression that the latter is related to security seeking, in contrast to Approach, which in its extreme form is nearer to recklessness). Similarly, women are more given to Flight than men, but there is no difference with respect to Fight. The fact that these components of the theoretical Activation and Fight/Flight systems are differentially distributed by sex raises further doubts about their unitary nature.

Figure 1 shows a configuration that appears reliable in that the male and female samples replicate one another, and this may be summarized in terms of three main clusterings. Approach and Fight go together in the general region of P and E from the Eysenck Personality Questionnaire. Passive Avoidance, Flight and Extinction are grouped with N, and Active Avoidance separates off towards the Lie (dissimulation) scale. Commonsense descriptions of these three main groupings might be 'courage', 'cowardice' and 'respectability'.

Conclusion

Somewhere between the observation of animal learning on which Gray's theory is based, and the organization of human personality revealed by use of the present questionnaire, relationships appear to have changed. The distance between these two levels of investigation is so great that there are many possible loci for the change (apart from the possibility that, even at the animal learning level from which it starts, Gray's theory is wrong). It is possible that, while we did our best to frame questions that would tap the same behavioural propensities as are measured in animal

experiments, we failed to do so, either because propensities of the appropriate kind (active avoidance, fight, flight, etc.) do not exist at the human level, or because they do not come out in answers to questionnaires as they might in laboratory or real-life measurements of behaviour. This possibility could be addressed by determining the degree to which our scales are able to predict such behavioural measurements.

Acknowledgement

We are grateful to Ruth Shelley for distribution of the questionnaires.

References

Boddy, J., Carver, A. & Rowley, K. (1986). Effects of positive and negative verbal reinforcement on performance as a function of extraversion–introversion: Some tests of Gray's theory. *Personality and Individual Differences*, **7**, 81–88.

Eysenck, H. J. & Eysenck, S. B. G. (1975). *Manual for the Eysenck Personality Questionnaire*. London: Hodder & Stoughton.

Gray, J. A. (1981). A critique of Eysenck's theory of personality. In H. J. Eysenck (Ed.), *A Model for Personality*. Berlin: Springer.

Gray, J. A. (1987*a*). The neuropsychology of emotion and personality. In S. D. Iversen, S. M. Stahl & E. C. Goodman (Eds), *Cognitive Neurochemistry*. Oxford: Oxford University Press.

Gray, J. A. (1987*b*). Perspectives on anxiety and impulsivity: A commentary. *Journal of Research in Personality*, **21**, 493–509.

Lewinsohn, P. M. (1975). The behavioral study and treatment of depression. In M. Herson (Ed.), *Progress in Behavior Modification*. New York: Academic Press.

Lingoes, J. C. (1973). *The Guttman–Lingoes Non-metric Program Series*. Ann Arbor, MT: Mathesis Press.

Nichols, S. & Newman, J. P. (1986). Effects of punishment on response latency in extraverts. *Journal of Personality and Social Psychology*, **50**, 624–630.

Patterson, C. M., Kosson, D. S. & Newman, J. P. (1987). Reaction to punishment, reflectivity and passive avoidance learning in extraverts. *Journal of Personality and Social Psychology*, **52**, 565–575.

Torrubia, R. & Tobena, A. (1984). A scale for the assessment of 'susceptibility to punishment' as a measure of anxiety: Preliminary results. *Personality and Individual Differences*, **5**, 371–375.

Wilson, G. D. (1975). The 'catchphrase' approach to attitude measurement. *Personality and Individual Differences*, **6**, 31–37.

Received 26 October 1988; revised version received 2 June 1989