

What People Will Do: Personality and Prediction *Peter Goldie*

I

People sometimes surprise us by the things that they do. What is the explanation of this phenomenon? This is a very general question, and I want here to treat it as such. I do, however, want to restrict the question to those surprising things that people do that are intentional actions, or things done for a reason in the particular sense that Elisabeth Anscombe was searching for in her book *Intention*.¹ For people sometimes do surprising things other than intentional actions. Talleyrand, the great diplomat who served the Ancien Régime, then the Revolutionary government, then Napoleon I, then the restoration of the monarchy with Louis XVIII, and then after the 1830 revolution with Louis-Philippe, was said never to do anything without intention, without a good reason. When he finally died in 1838, Metternich, surprised, made this famous remark: 'I wonder what he meant by doing that'.

People don't die intentionally, for reasons in the sense required for meaningful intentional action (suicide is something quite different). And that's what makes Metternich's remark into a nice joke, as well as indicating what is supposed to be special about action as such, according to which dying is not a kind of action.

So let me from now on limit my question to intentional action, which I will characterise as something done for a reason, consisting of a belief and an attitude towards this kind of action (let's from now on call it a desire, in the knowledge that this term is desperately vague; but nothing hangs on it here). And this belief and desire will give the answer why, in the sense we want, people do the things that they do; the belief and the desire rationalise the action.² Someone opens the door of the fridge, and he does this because he wants a beer and believes that opening the door is the best means of getting a beer, believing, as he does, that there is a beer in the fridge. Someone takes an umbrella on her walk, and she does this because she believes it's going to rain later on, because she wants not to get wet, and because she believes that taking an umbrella is the best means of avoiding getting wet. This kind of means-end, belief-desire explanation, it is said, can be extended to explain much more complex examples of intentional action.

I accept that, in respect of all intentional actions, these kinds of belief-desire explanations are available. Indeed, it seems to me that, in our thinking about action, our own and other people's, we just *take it for granted*-it is, I dare say, a priori-that belief-desire explanations are always going to be available; although, of course, they may not always be found.

However, a moment's reflection on these belief-desire explanations should reveal just how thin they are, partly because, in turn, they rely on such a thin notion of rationality. To rationalise an action in this thin sense is just to show how it could make sense for someone to do such a thing. And it can often make sense to do one of a diverse range of possible things. In a restaurant you are brought the wrong flavour of ice-cream: you ordered chocolate and you are brought strawberry. It would make sense to tell the waiter he has made a mistake; it would make sense to eat what you

are given; it would make sense to leave it; it would make sense to throw the icecream on the floor and walk out of the restaurant; it would make sense to offer the ice-cream to your partner; it would make sense to pour the icecream onto your partner's lap; and so on. If you do one of these things, it's very likely that you didn't at the time consider all the other possible courses of action that were open to you, or perhaps any of them; the other possibilities might not have entered your mind. So we must not forget that a belief-desire explanation doesn't explain why someone did one thing *rather than another which was also open to him*.

The point emerges most starkly when we turn to prediction: I ask you to predict what a rational person (that's all we know about him or her) will do if he or she is brought the wrong flavour of ice-cream by a waiter in a restaurant. 'Well,' you will sensibly reply, 'it depends.'

One of the things it depends on is personality. Because I know my wife's personality, I can pretty reliably, I think, predict what she would do in such circumstances. However, I acknowledge that my prediction might be wrong: she might on a particular occasion surprise me and do something other than what would be expressive of her personality. So what she does must depend on more than just her personality.

To explain why I am right to acknowledge that my wife might surprise me in what she does, I need to turn to E. M. Forster's distinction, which he makes in his *Aspects of the Novel*, between flat and round characters. 'Flat characters', he says, 'were called "humours" in the seventeenth century, and are sometimes called types, and sometimes caricatures. In their purest form, they are constructed round a single idea or quality ... The really flat character,' he continues, 'can be expressed in one sentence such as "I will never desert Mr Micawber". There is Mrs Micawber-she says she won't desert Mr Micawber; she doesn't, and there she is'.³ Flat characters are contrasted with round characters. 'The test of a round character,' Forster says, 'is whether it is capable of surprising in a convincing way. If it never surprises, it is flat. If it does not convince, it is flat pretending to be round. It [the round character] has the incalculability of life about it'.⁴

It is because flat characters have this 'single idea or quality' that they never surprise us; everything they do is determined by whatever quality or characteristic they are deemed to have, and they never act contrary to, or against, their type. We real life human beings, in contrast, like my wife, are round characters-all of us. So, by definition, we round characters are capable of surprising, and when we surprise, we do so in a convincing way.

The fact that people can surprise us by what they do shows that we expected them to do one particular thing, or one of a range of things, and in fact they did something different. Expectation in this context is, I think, something less than prediction, although it can extend to prediction. Expectation can involve simply taking some action or other for granted, rather as you expect-take it for granted-that someone who buys a cup of coffee in Caffé Nero will drink it and not throw it on the floor, or that a tennis player in the middle of a game will try to return the ball and not try to imitate Marlon Brando. Expecting what someone will do in some circumstance or other, as my ice-cream example shows, and as these examples show, involves more than just

expecting what any rational person would do: the idea of rationality has to be thickened out with *personality*. So, given a particular person's personality, my wife's for example, we can expect that she will do one kind of action amongst a range of possible rational actions, and expect that she will not do other kinds of action which it would be equally rational to do. For someone else with a different personality from my wife's, we may expect them to do something else, and not what we expect my wife to do. This, personality and expectation or prediction, is the first thing I will try to explain.

The second thing I will try to explain is the confounding of our expectations and predictions-the fact that we can be surprised, in a convincing way, when people do something other than what we expect-when they act contrary to what their personality traits lead us to expect.

Then, thirdly, I will try to draw some practical conclusions from all thisthings that we round characters might bear in mind in leading our lives, and in judging ourselves and others. But let me begin by saying something about what I mean by personality

II

What is a personality trait? Personality traits-and here I include character traits in this category-are dispositions. All sorts of things have dispositions. This glass is fragile. Its being fragile (its fragility) is a disposition of the glass. Dispositions like these can be understood in terms of 'if-then' conditional statements: 'if this glass were to be thrown against the wall, then it would break'.

When we come to personality traits, to say what the 'if-then' conditional looks like *in general* is, I think, a hopeless task, because personality traits are so disparate in kind. There are ways of acting, such as being charming; there are relatively enduring temperaments, such as being cheerful, being nervous and being gloomy; there are relatively enduring emotional dispositions, such as irascibility and being envious; there are relatively enduring preferences and values, such as being a book-lover, being a foodie, liking football, and disliking authority figures; there are talents, such as being quick-witted; and there are character traits, which are, roughly, relatively enduring dispositions to have certain kinds of motives in certain kinds of situation, and thus to act in certain kinds of ways.

'If-then' conditional statements for personality traits are what Gilbert Ryle has called 'inference-tickets'-'season tickets' which allow us to explain and predict motive and action on a particular occasion.⁵ For example, knowing that someone is a polite and considerate person allows us to predict that she will have particular polite and considerate motives if she sees that her fellow dinner-guest needs the salt, and thus that she will pass the salt in a polite way.

Flat characters have a single personality trait, which can be utterly relied upon to issue in the relevant kind of action on all occasions. Thus flat characters do not-they cannot surprise us. But, as Forster rightly says, flat characters are only to be found in novels; it is a kind of prejudice-a kind of stereotyping-to flatten out real life people. So, for us real life people, we need to build into our notion of personality this capability of surprising us. Any notion of reliability, of predictability, in motive and action must

allow for the possibility of a failure to be motivated or to act as our ascription of personality traits allows us to infer. One might put it like this: an inference-ticket, a season ticket, can fail on an occasion to get you to your expected destination.

We could just build in to the 'if-then' conditional a 'normally', a 'usually', or an 'all other things being equal', and this would certainly allow for the possibility of being surprised: a polite person will *normally* or *usually* have polite motives and act politely when it is appropriate. Fair enough. However, although this might explain how we can be surprised-on those occasions when the 'normal' or 'usual' doesn't apply-it isn't going to be enough to explain how we can be surprised *in a convincing way*. So let's now look at the kinds of thing that can provide convincing explanations of our surprising failures to do as our personality traits imply that we should.

III

To remind you of what I said earlier, I accept that all intentional action can be explained by appeal to beliefs and desires of the individual. However, belief-desire explanations, thin as they are, can be thickened out by appeal to factors that influence the way someone's mind works on a particular occasion, which can help to explain why someone has motives-beliefs and desires, intentions and so on-which are not what we would expect from his personality, and which, in turn, can help to explain his acting contrary to how we expected, and in a way that initially surprises us. These factors are not themselves entirely within what has been called the 'space of reasons': they so to speak bridge the divide between the mental and the physical; or rather, they throw into question the very idea of there being such a divide. I will put them into four broad categories.

First, there are states such as being drunk, being under the influences of drugs, having a bad cold, and being deprived of sleep.⁶ A shy and retiring man goes to a cocktail party. He bumps into a woman whom he hasn't seen for many years. Suddenly, he is struck by the thought that he has been in love with her all this time, and, quite out of character, he tells her as much. We explain this by saying that he had had two glasses of champagne on an empty stomach; it was because he was a little bit drunk that he did what he did. What we would expect him to do, given his personality, is to have characteristically shy and retiring thoughts-'I couldn't possibly force my attentions on her; I hardly even know her.' Perhaps on this occasion he did have these thoughts, but normally, usually, we would expect them to have prevailed over any amorous inclinations he might have felt; and yet on this occasion they didn't prevail. You might think that this kind of case exemplifies weakness of the will. But weakness of the will is something of a moral notion: it is not straightforwardly obvious that the champagne *distorted* his thinking, and led him to allow *inappropriate* deliberative influence to his amorous inclinations: thinking and acting out of character because of drink and having your thinking distorted because of drink are not necessarily the same thing. Maybe she will turn out really to be the love of his life, in which case, in retrospect, we might come to agree that, without the champagne, being such a shy and retiring person, he would have missed out on the chance of his life. The champagne helped him to relax, to properly lift up his eyes to her beauty and her talents, and to enable him to see her as she really is. As Iris Murdoch put it, 'By opening our eyes we do not necessarily see what confronts us. We are anxiety-ridden animals. Our minds

are continually active, fabricating an anxious, usually self-preoccupied, often falsifying veil which partially conceals the world.’⁷

My second category is emotions like being angry and being jealous, which, although capable of being justified or grounded in reasons, can influence thinking in ways that lead people to act contrary to what we would otherwise expect and in an unjustified way. You shout at your child for not sitting up straight in his high chair, in spite of being a caring and loving parent. You do this because you are angry. Your shouting at her can also be given a belief-desire explanation: you wanted her to sit up straight, and you believed that shouting ‘Sit up straight!’ at her was the best means of getting her to do this. But why did you, a caring and loving parent, have these thoughts, so unusual for someone like you? Because you were unjustifiably angry, and because shouting at people who won’t do what you want is characteristic of angry behaviour.

My third category is moods, relatively short-term states like being depressed, tense, irritable, full of unlocated sexual desire. Her not going to work today, in spite of being a diligent and hardworking person, and much to the surprise of her bosses, who has expected her to be there, can be explained by appeal to the fact that she felt depressed when she woke up this morning. Her being in this mood-depressed- helps to explain why she had the thoughts that she did, and thus goes beyond the belief-desire explanation. Perhaps what went on in her mind at the time was the thought that she might not be able to handle this large and important deal, and the thought that if she didn’t turn up to work, then someone else would take the deal off her hands. Our thicker explanation-she was depressed-points towards an explanation of why she had those thoughts in spite of being normally so diligent and hardworking. Moods, in turn, can explain emotions, and they often do so without justifying them: for example, you got unjustifiably angry with your child because you were tense and irritable.

Sometimes, and this is my fourth category, the explanation of an action that surprises us appeals to the influences of the particular situation that the individual finds himself in: his being in that situation brings about a surprising influence on thinking. The literature in social psychology these days is replete with examples which bring into question the reliability of personality traits in certain situations.

I will mention just one. In the summer of 1971, Philip Zimbardo, a psychologist at Stanford University, carried out an experiment on the effect of imprisonment. The account that follows is either direct citation from Zimbardo or paraphrase. He and his colleagues took a sample of 24 college students from the U.S. and Canada who happened to be in the Stanford area and wanted to earn \$15 a day by participating in a study. On all measured dimensions, these students were typically healthy, intelligent, middle-class males. The participants were arbitrarily divided into two groups by a flip of the coin. Half were randomly assigned to be guards, the other to be prisoners.

The ‘prisoners’ were brought to the ‘jail’ (in fact the converted basement of the psychology department) one at a time and greeted by the warden, who conveyed the seriousness of their offence and their new status as prisoners. Each prisoner was systematically searched, stripped naked, and then deloused and issued with a uniform, a heavy chain bolted on the right ankle and worn at all times, rubber sandals, and a stocking cap made from a woman’s nylon stocking. The guards were given no specific training on how to be guards. Instead they were told that they were free,

within limits, to do whatever they thought was necessary to maintain law and order in the prison and to command the respect of the prisoners. As with real prisoners, the prisoners in this experiment expected some harassment, to have their privacy and some of their other civil rights violated while they were in prison, and to get a minimally adequate diet—all part of their informed consent agreement when they volunteered.

Because the first day passed without incident, Zimbardo and his colleagues were surprised by, and totally unprepared for, the rebellion which broke out on the morning of the second day. The prisoners removed their stocking caps, ripped off their numbers, and barricaded themselves inside the cells by putting their beds against the door, taunting and cursing the guards. The guards met force with force: they used fire extinguishers to get the prisoners away from the doors; they broke into each cell, stripped the prisoners naked, took the beds out, forced the ringleaders of the prisoner rebellion into solitary confinement, and generally began to harass and intimidate the prisoners. Once the riot was under control, they stepped up the violence and intimidation techniques. Every aspect of the prisoners' behaviour fell under the total and arbitrary control of the guards. Even going to the toilet became a privilege. Prisoners were often forced to urinate or defecate in a bucket that was left in their cell, and sometimes the guards would not allow prisoners to empty these buckets, so that the prison began to smell of urine and faeces. After a rumoured prison break, which in fact turned out to be false, the guards again escalated very noticeably their level of harassment, increasing the humiliation they made the prisoners suffer, forcing them to do menial, repetitive work such as cleaning out toilet bowls with their bare hands.

After five days, prisoners were seen being marched on a toilet run, bags over their heads, legs chained together, hands on each other's shoulders. At this point, long before the scheduled end of the experiment, Zimbardo was made to realise the psychological damage that the experiment was causing to the participants, and he quickly brought it to a close. To quote Zimbardo, 'We had created an overwhelmingly powerful situation—a situation in which prisoners were withdrawing and behaving in pathological ways, and in which some of the guards were behaving sadistically. Even the "good" guards felt helpless to intervene, and none of the guards quit while the study was in progress'; '...not once did any of the so-called good guards ever contest an order by a sadistic guard, intervene to stop or prevent despicable behaviour by another guard, or come to work late or leave early'. He reports that, amongst both guards and prisoners, '[N]one of our preliminary personality tests were able to predict this behavior. The only link between personality and prison behavior was a finding that prisoners with a high degree of authoritarianism endured our authoritarian prison environment longer than did other prisoners.'⁸

Given this dissociation of behaviour and personality, it would be a mistake to make an inference from a given type of behaviour on this single occasion, such as the brutal behaviour of many of the guards, to a brutal personality trait, and then appeal to that trait to predict and retrodict further brutal behaviour across a range of different kinds of situation. It isn't at all clear how to explain what these people did in this prison experiment, but I have no doubt that a simple 'brutal is who brutal does' is far too quick. The many experiments in social psychology show both how mistaken this is, and how prone we are to make the mistake.

Being drunk, being angry, being depressed, becoming a prison guard in particular circumstances: a large part of our everyday psychology is concerned with thicker explanations that appeal to factors such as these. And factors from each of my four rough categories can interweave in a complex network of narrative explanation: emotion, such as anger, can be explained by mood, such as being tense and irritable, as we saw earlier; being tense and irritable can be explained by lack of sleep and by situation, such as being harassed by a domineering boss; and so on. These narrative explanations explain through showing, sideways on so to speak, why people's patterns of thinking-their occurrent thoughts and choices-are as they are, and why they come to do things which surprise us. We find out, through our own experience, through literature and the arts, through reading newspapers, what characteristic influences various kinds of factors have. Being drunk often explains why people give way to temptation; we all know it can be a terribly expensive thing to go shopping after a boozy lunch. Being angry often explains why people do things like shout at their child or throw the wrong flavour of ice-cream on the floor; we all know the best thing to do when you're angry is to count backwards from a hundred. Being depressed often explains why we shirk from getting down to our work; we all know this too. Being thrown into an unusual situation without proper preparation, such as becoming a prison guard, can often explain why you behave as you shouldn't. It's all very human and understandable.

IV

What practical conclusions can be drawn from all this? I would like to mention two, both of which bear specifically on character, rather than more widely on personality. The first concerns judgement. Let's say that some influencing factor leads you to act out of character in ways that were unexpected, and that this involved acting badly. We all thought, as did you, that you had some virtue-courage, say-and yet when you come to be tested you failed to act as you should have done. At the crucial moment, you acted to save your own life, in disregard of the others whom you should have helped.

Now, what I've been saying suggests two things. First, and this I mentioned earlier in relation to the Zimbardo prison experiment, we shouldn't simply brand you as having a bad character, as being a coward, 'cowardly is as cowardly does', just because of this one action. It's quite possible that you really are a courageous person, but less than fully virtuous, so courageous action isn't guaranteed on every occasion. After all, only flat characters are fully virtuous, and you are a round character. Secondly, we are round characters too, so we cannot rest complacently in the secure knowledge that our own virtue will smoothly generate virtuous action in all circumstances. So when we hear of what you did-that so-called cowardly action-we ought to reflect on the fragility of character, and acknowledge that we too might have done what you did if we had been in your circumstances. For who *really* knows what we might have done? As Marlow, Joseph Conrad's narrator in *Lord Jim*, said, 'Let no soul know, for the truth can be wrung out of us only by some cruel, little, awful, catastrophe'.⁹

However, it is consistent with our thinking that we might have done the same as you in those circumstances, or even that just about anyone would have done the same as you in those circumstances, also for us to think that what you did was wrong, that you are responsible for what you did, and that you are to blame for it. Similarly, you should blame yourself for what you did. You should not hide behind the influencing

factor-being a participant in the prison experiment for example-as an excuse to get you off the hook; as J. L. Austin said, some excuses get us on the hook rather than off it.¹⁰

But-and this is a point I cannot argue for here-your blame of yourself should not be misdirected, in remorsefully and remorselessly-*looking backwards*, back to that one moment in your life, that one moment of failure of character. Without dodging responsibility, and without avoiding blaming yourself, you can focus your thoughts not, remorsefully, on the past, and on what you did wrong at that single, fateful moment, but on the future, and on changing yourself so that you won't make the same mistake again. As Nietzsche so brilliantly put it: 'Never yield to remorse, but at once tell yourself: remorse would simply mean adding to the first act of stupidity a second. - If we have done harm we should give thought to how we can do good'.¹¹

The second practical conclusion, also bearing on the fragility of character, is that we should be circumspect about the reliability of our character and our motives on any particular occasion. The real springs of human action are a mystery, as much our own actions as those of others; and sometimes more so. Because we have names for things- 'motive', 'deciding', 'willing', and so on-we too readily conclude that there is something very clear and precise that the names stand for. We may be able to deliberate about our motives, decide what to do, and later explain or make sense of what we have done-using names for motives, for deciding, for willing. But still, to quote Nietzsche again:

... at the moment when we finally do act, our action is often enough determined by a different species of motives than the species here under discussion, those involved in our 'picture of the consequences'. What here comes into play is the way we habitually expend our energy; or some slight instigation from a person whom we fear or honour or love; or our indolence, which prefers to do what lies closest to hand; or an excitation of our imagination brought about at the decisive moment by some immediate, very trivial event; quite incalculable physical influences come into play; caprice and waywardness come into play; some emotion or other happens quite by chance to leap forth; in short, there come into play motives in part unknown to us, in part known very ill, which we can *never* take account of *beforehand*.¹²

One might almost think that Nietzsche had in mind the recent experiments in social psychology when he wrote that. The idea, then, is this: if we can't be sure of our own motives on an occasion, or of our character, then we would do well to realise this in advance of the moment of action, and to plan accordingly. This is what I mean by circumspection-as the OED has it, 'attention to circumstances that may affect an action or decision; caution, care, heedfulness'. Like Odysseus, who had his crew tie him to the mast so he couldn't sail towards the tempting Sirens, we should sometimes plot against our future selves, by putting things in the way to prevent our being unduly influenced to act out of character.¹³ Circumspection in forward planning is, like strength of will, a kind of executive virtue. But it comes temporally prior to strength of will; without the proper circumspection about our motives and character, strength of will can both be not enough and come too late. Don't rely on your character or your strength of will to see you through when you are tempted; better to keep yourself out of the way of temptation, or to keep temptation out of your way.

There's an interesting implication of this. If you acknowledge, as I say you should, that you are less than fully virtuous, not wholly and completely reliable in your actions, even (in that telling expression) with the best will in the world, then, somewhat ironically, it is a mistake to ask yourself, in thinking about what is the right thing to do, 'What would the virtuous person do here?'. For the *really* virtuous person, if there were such a flat character outside the pages of a novel, wouldn't be led astray, and could therefore *blithely* go into these situations, in which the rest of us would be tempted. So we shouldn't do what the fully virtuous person would do. The better question to ask is, 'What would the virtuous person *advise* me to do here?'. The virtuous person may well wisely advise you to be circumspect about your character and motives, and, like Odysseus, to take steps in advance to make it harder for you to act out of character in ways that you may well later come to regret.¹⁴ We are all round characters, capable of surprising in a convincing way. Accordingly, we must not, like Conrad's Marlow, hope for the impossible, 'for the laying of what is the most obstinate ghost of man's creation, [for the laying] of the uneasy doubt uprising like a mist, secret and gnawing like a worm, and more chilling than the certitude of death the doubt of the sovereign power enthroned in a fixed standard of conduct'¹⁵

Peter Goldie
King's College London

¹ *Intention*, Oxford: Blackwell, 1958.

² See especially Donald Davidson's 'Actions, Reasons, and Causes', together with the other papers in his *Essays on Actions and Events*, Oxford: Clarendon Press, 1980.

³ *Aspects of the Novel*, London: Pelican Books, 1962, p. 73.

⁴ *Aspects of the Novel*, p. 81.

⁵ *The Concept of Mind*, London: Penguin, 1990, p. 117; first published by Hutchinson, 1949.

⁶ These are called 'non-rational influences on thinking' by Jane Heal in 'Replication and functionalism,' in M. Davies and T. Stone, eds., *Folk Psychology: The Theory of Mind Debate*, Oxford, Blackwell, 1995, pp. 45-59. I discuss these issues in *The Emotions: A Philosophical Exploration*, Oxford: Clarendon Press, 2000, pp. 167-75.

⁷ Iris Murdoch, *The Sovereignty of Good*, London: Routledge, 1970, p. 84.

⁸ Citations from <http://www.prisonexp.org/>, and from 'Reflections on the Stanford Prison Experiment: Genesis, Transformations, Consequences', Zimbardo, P. G., Maslach, C., & Haney, C., in T. Blass, ed., *Obedience to authority: Current perspectives on the Milgram paradigm*, Mahwah, NJ: Lawrence Erlbaum, 2000, pp. 193- 237. See the website for more references.

⁹ *Lord Jim, A Tale*, Oxford World Classics, Oxford: Oxford University Press, 2002, p. 236.

¹⁰ J. L. Austin, 'A plea for excuses,' reprinted in his *Philosophical Papers*, 3rd edition, Oxford: Oxford University Press, 1979, pp. 175-204.

¹¹ In his *Human All Too Human*, 'The Wanderer and his Shadow', Section 323, translated by R. J. Hollingdale, Cambridge: Cambridge University Press, 1986.

¹² *Daybreak*, translated by R. J. Hollingdale, Cambridge: Cambridge University Press, 1982, Section 129.

¹³ There is an interesting book about all this: *Ulysses and the Sirens: Studies in Rationality and Irrationality*, by Jon Elster, Cambridge: Cambridge University Press, 1979.

¹⁴ For discussion of the 'advice' and the 'example' or 'emulation' models, see Michael Smith, *Philosophy and Phenomenological Research* 55, 1995, pp. 109-31. The question people sometimes ask is 'What would Jesus do?', and it's interesting in this context to note that Jesus was tempted.

¹⁵ *Lord Jim*, p. 37.