Ethical Thought after Situationism

Pauline is a woman in her late 30s, a married mother of three.

1. She is regularly kind to her children. She often sacrifices the pursuit of her own ends so that they can attend various practices, go on various trips and socialize with their friends.

2. However, she finds herself mostly unable to make any similar sacrifices for her husband, as their marriage has become somewhat lifeless. Nonetheless, on certain special occasions, she surprises him with acts of kindness.

3. She is known for being compassionate towards her friends, but most of her friends know that when her marriage or her career are causing difficulties for her, she is likely to ‘lash out’.

4. Finally, she is a ruthless careerist in her workplace, advancing her status in the organization with little to no regard for the needs or feelings of her co-workers.

Is Pauline a compassionate person? It certainly doesn’t seem so, as she fails to be consistently compassionate over a wide range of contexts. Is she, then, only moderately compassionate? Or is the character-concept ‘compassion’ too coarse-grained to apply to someone like Pauline?

John Doris thinks that nearly all of us are basically like Pauline in one very important respect. He has marshalled a mountain of social-psychological data which purports to show that human behaviour is massively situational. In other words, our behaviour, like Pauline’s behaviour, is much better explained by situational factors than by any alleged character trait or stable disposition. Doris alleges that this refutes globalism, or the view that there are situationally consistent, temporally stable character traits. I find Doris’ anti-globalist arguments convincing, and I will offer only a cursory discussion and defense of them in this paper.

Doris, however, embraces the possibility of local character traits:

In my interpretation, situationism does not entail an unqualified skepticism about the personological determinants of behavior; I allow for the possibility of temporally stable, situation-particular, "local" traits that are associated
with important individual differences in behavior. As I understand things, these local traits are likely to be extremely fine-grained. ¹

I will argue, contra Doris, that we should not subscribe to the notion of a local character trait, and that this implies the empirical inadequacy of all characterological ethical discourse. We should, I argue, focus on the psychological basis for such context-sensitive behaviour, rather than posit complex dispositional traits. While this conclusion appears initially troubling, I will further argue that the private trust we place in our closest friends and companions is not underwritten by any ‘local’ disposition or trait, but rather by the particular roles they play in our lives and by associated mental states like love and care. Relatedly, I hold that the public trust we normally place in our fellow citizens is justified by their identification with a social role and its corresponding norms and values. In saying all of this, I mean to imply that role-based evaluation is a powerful and important feature of ethical life, and that the failure of disposition-based ethical evaluation is apt to make us far less anxious when we maintain our grip on these facts.

Against Local Character Traits

According to Doris, global character traits do far less explanatory and predictive work than then they need to in order to count as a viable evaluative descriptors. Consider the trait “compassionate”. If Doris is right, and situationist experiments show that almost everyone is like Pauline, then this trait can only play a very restricted role in an explanation of how persons act in various situations. Furthermore, given that our behaviour is apparently very sensitive to the most trivial of situational factors, we cannot reliably predict whether or not the average person will act compassionately in a given situation, even with a significant amount of information about their past behaviour. For these reasons, the evaluative role of the disposition-concept “compassion”

appears to be significantly deflated. It might have been thought—and it certainly has been thought by Aristotelean moral theorists—that the concept could be meaningfully applied in everyday ethical situations. As it turns out, the large majority of persons are not distinguishable in terms of their compassionate behaviour. As such, the idea of “compassion” can only, at best, mark out certain comparatively rare characters whose behaviour we might strive to emulate, even if most of us will usually fail to do so.² Given that a fully virtuous person is not only compassionate, but also honest, courageous and perceptive, it does appear that fully virtuous characters will be extremely rare.

Now, recall Doris’ claim that we might still identify and deploy local character traits. These are dispositions that have specific situational factors built directly into their description. Pauline, for example, is not compassionate, rather, she is, to some degree,

\[\text{L1}\text{ “compassionate-towards-friends-when-not-under-work-related-stress-or-marital-stress”}.\]

Knowing what we know about Pauline’s behavioural patterns, we can immediately see that Doris is right to say that “there is a markedly above chance probability that the trait-relevant behavior will be displayed in the trait-relevant eliciting conditions.”³ However, the scope of local traits is greatly restricted, both in terms of the set of individuals to which they can apply and in terms of the number of situations in which they can apply. The more fine-grained a local trait becomes, the more narrow its scope will be in these respects. Does this matter?

Given that Doris thinks that empirically adequate local traits will be “extremely” fine-grained, I believe that it does. Imagine if ceramic vases only tended to shatter when dropped from very specific heights, say, three, six and eighteen feet. In this case, the global disposition \textit{fragility}

---

² It might be suggested that this is all character-concepts ought to do. Doris (1998) anticipates and responds to this challenge.
would certainly become explanatorily and predictively useless in reference to vases, in just the same way that the trait ‘compassion’ is not useful for persons. It seems that we would be presented with two choices. We could invent three separate “local” dispositions and ascribe them to ceramic vases. Alternately, we could abandon the search for dispositional properties and search for the underlying causal story provided by what some metaphysicians have called the explanatory ground, or the complex set of facts about the object that explain its irregular behaviour.\(^4\) In other words, some basic set of facts about the object must explain why it only shatters when dropped from these very specific heights, and we would normally attempt to replace the fine-grained dispositional property with an explanatory story featuring that set of facts.

In the case of local character traits, I believe that this is precisely what we should do. Because of their greatly reduced scope, it is not clear how local traits can play a significant explanatory, predictive or evaluative role in ethical discourse. Absent some description of how they might be used to talk about persons, the case for such traits seems rather thin. Furthermore, in the case of supposed “dispositions” like L1, we can replace the fine-grained local trait with a general description of the mental states that lead to the behavioural pattern in question. Given the failure of global character traits, I am suggesting that we return to the relevant ‘explanatory ground’ and try to characterize human behaviour in terms of the mental states that produce it. In Pauline’s case, for example, we should replace local trait L1 with a psychological explanation: she cares about her friends, but when her project-related desires are frustrated, this care is much less motivationally effective. Furthermore, rather than say that she is “compassionate towards one’s children”, we should say that she loves her children. These explanations can be integrated into a more general psychological picture of Pauline, whereas it is unclear how local traits might

help to comprise such a theory. Doris, to his credit, recognizes that a “more enlightening”
description of local traits would make reference what he calls “the psychological context of the
trait”5, but I am suggesting that the alleged trait is its psychological context. I think that Doris’
position entails the outright reduction of character-discourse to mental-state discourse, not its
retention in some needlessly limited form.

Is there nothing, then, in the idea of character? Robert Solomon has given voice to a deep
and troubling anxiety surrounding the situationist’s contention that there are no global character
traits:

Do you trust your wife? Your husband? Your friend? Your colleague? Your editor? Your students? On what basis does anyone trust anyone? Do we really trust them? But for trust or knowledge, the presumption is that there is a basis for this in the person, a stability or regularity of character… I trust my wife because I know her. I recognize in her a character that is robust and resolute.6

Doris is right to say that any set of ethical ideas that emerges from the shadow of Situationism
must be ‘empirically adequate’. However, Solomon is alerting us to a second requirement: we
must be able to make sense of our actual trust in our private and public relationships, for as he
points out, this trust is both pervasive and non-optional. I have so far tried to suggest that
Solomon’s trust in his wife, like the trust Pauline’s children might place in her, need not be
grounded in dispositional regularity. Rather, it is a trust that is grounded in love or care, in
particular mental states that need not display any special kind of generality. I will now proceed to
argue that that we have context-sensitive materials with which to explain, predict and evaluate
human behaviour. These materials are provided by social roles.

5 (Doris 2002) p. 66.
Schemata and Social Roles

By way of introduction to this topic, I want to consider a social-psychological study which has not received much attention amongst philosophers working on situationism. In 2004, Varda Liberman, Lee Ross and Steven Samuels conducted two separate experiments in Stanford, California and Herzliya, Israel. Subjects were given a standard prisoner’s dilemma game to play, with a fairly simple cost-benefit matrix and a pre-defined set of iterations. This experiment, however, had its own peculiar twist: the subjects were divided into two separate groups, and while one group of subjects were casually informed that the game was called “Community”, the other group was informed that the very same game was called “Wall Street”. The results were extraordinary: co-operation in “Community” outstripped co-operation in “Wall Street” by a factor of 2:1 (70% to 35% co-operation, roughly). Perhaps most striking is the fact that in the announced final round of the prisoner’s dilemma, when it is always subjectively irrational for a player to co-operate, “Wall Street” players almost universally defected, while “Community” players were only slightly more likely to defect. This was in spite of the fact that the subjects believed that they were playing for substantial amounts of money.7

Now, on the face of it, we might think that this study is simply one more bullet in the situationist’s ammunition-box, another strong piece of evidence for the conclusion that situations regularly overwhelm persons. Jesse Prinz appears to draw this inference, claiming that “in line with situationist psychology, the label itself had a dramatic impact…[a] situational variable overwhelms our capacity to make decisions.”8 Prinz, like many philosophers drawn to situationism, prefers an external explanation of subject behaviour, claiming that it is “the label”

8 Prinz 2009, p.140. Prinz also dramatically overstates the results of the study, claiming that the rate of co-operation was four times higher in the community-game. While I’m sure the mistake was unintentional, it is the sort of error which can give Prinz’s readers the impression that things are much simpler than they actually are.
(and not any person’s perception of or interaction with the label) that is ultimately responsible for the experimental results. This is odd, because if a bodily movement is explained only by factors external to the body in question, then we might think that, under this description, the movement is not an action. In other words, if bare situational elements are the only factors that explain a person’s bodily movement, then the movement is probably not ethically significant, nor is it the proper object of study for psychology. Internal explanations make reference to the subject’s frame of mind, her perception of the situation, and so forth. The extent to which we provide internal explanations of situationist results is the extent to which we are studying behavioural psychology, and an internal explanation is just what Liberman, Ross and Samuels provide. They claim that their results are best explained by the fact that the subjects were primed to adhere to behavioural norms associated with a role-stereotype:

These findings in no way negate the fact that different individuals may interpret particular types of situations differently… Rather, the findings of our present research suggest that most of us are equipped with a range of different cognitive schemata that could be applied to a given situation and that our behavior may depend on whatever factors determine which schema happens to be recruited by chance or by subtle features of the context at hand.

The defender of local trait-attribution might want to claim, rather unhelpfully, that people are “Wall-Street-Game-Uncooperative” and “Community-Game-Cooperative” to varying degrees. However, the study’s authors wisely avoid the search for a dispositional interpretation of the results. Instead, they produce an internal explanation by making reference to what I have been calling the explanatory ground: the psychological states that produced the behaviour in question. We might augment their brief remarks with reference to what social-cognition theorists call a role schema. Bailey and Yost (2000) provide this helpful definition:

A schema is a highly ordered cognitive structure composed of knowledge, beliefs, and feelings about persons, objects, and events. Role schemas are those that organize proper behavioral patterns according to perceived social position.\(^{11}\)

A role-schema is a mental construct that is activated when a person perceives that either they or another person are occupying a social role. For our purposes, we can say that a role schema consists in a set of behavioural norms along with certain characteristic goals or aims strongly associated with the role in question. Many social-psychological studies have shown that the activation of a role-schema can dramatically affect an individual’s perception of what kinds of behaviour are appropriate, acceptable or desirable.\(^{12}\) Most of us believe (correctly, it seems) that Wall Street is the sort of place where individuals will only co-operate if it is in their immediate advantage to do so, and that successful brokers are those who steadfastly conform to this competitive standard. To be told that a game is called “Wall Street” is to have this cognitive schema and the associated behavioural norms ‘activated’. Furthermore, when schemata are activated in group situations, we can observe a feedback-loop, as the role-stereotypes are progressively entrenched and come to exert a kind of automatic motivational authority. This explains, in particular, why individuals in the “Community”-game did not defect in the final round. The motivational force of a substantial monetary reward was outweighed by the importance the community-member-scheme had acquired through successive rounds.

This schema-centred interpretation can help us to make sense of Phillip Zimbardo’s Stanford Prison Experiment. Each of the subjects in this experiment was given one of two labels, and while “guards” were not instructed to be cruel towards “prisoners”, they famously took the

---

initiative and devised countless methods of aggressive brutality. Christina Maslach writes that over the course of the experiment, prisoners and guards “all internalized a set of destructive prison values that distanced them from their own humanitarian values.”\textsuperscript{13} I find this picture of previously sensitive and compassionate subjects being unexpectedly invaded by a set of alien values highly implausible. The role-schema model suggests that that each subject came to the experiment with a previously established prisoner-schema and guard-schema, and it seems that both schemata were activated by various situational cues.

Because role-schemata can wreak such havoc, we might be tempted to think of them as intrinsically problematic. However, I think that this is precisely the wrong conclusion to draw. For example, we might be less inclined to make such a judgment about the Liberman, Samuels and Ross experiment, where the community-schema provoked co-operative behaviour. Schemata cannot be intrinsically troubling from an ethical point of view, rather, their activation can be more or less appropriate depending on the context.

Furthermore, we might be tempted to think that role-schemata undermine autonomy. Doris himself thinks that the Stanford Prison Experiment shows that people have less control over themselves than we might have thought. “Once again,” he writes, “it appears that persons are swamped by situations.”\textsuperscript{14} This resonates with Zimbardo’s contention that his subjects became both “dehumanized” and “de-individuated.”\textsuperscript{15} Nonetheless, I think we should resist the judgment that an activated role-schema necessarily undermines a person’s autonomy, or that a person ‘disappears’ when they conform to a stereotype. Observing subjects being made into monsters or puppets by an experimental setup may provoke this feeling, but we should not allow

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{13} Zimbardo, Malasch and Haney, 2002, p.18.
\item \textsuperscript{14} Doris 52.
\end{itemize}
\end{footnotesize}
the strangeness of experimental situations to infect our judgments about autonomy as it functions in reality.

In order to see why this is so, we need to return to Pauline. Suppose we observe her performing some self-sacrificing act for her child: she drives her daughter to swim practise at 5:30 AM, even though she knows that being sleep-deprived will impede her performance at work later that day. We ask her, incredulously, how she is able to do such things.

Suppose she says this: “I’m a mom. That’s just what moms do.” Pauline would be telling us that she is committed to adhering to the behavioural standards supplied by her mother-schema, even if she sometimes experiences strong desires to do otherwise. In this case, it seems, we might not be so sure that she is non-autonomous, even though she is completely under the sway of a role-stereotype, just as the Stanford Prison subjects were. One relevant difference in this case is clear; to use Harry Frankfurt’s influential terminology, she identifies with her role. Indeed, her final answer to our line of questioning is a kind of performative demonstration of the wholehearted, higher-order, reflective endorsement characteristic of Frankfurrian identification. As such, it makes no sense to question the schema’s authority over her will, for her will is partly constituted by the schema itself; as she has already told us, it’s what she is. “Individuals,” writes Frankfurt, “are autonomous to the extent that they govern themselves… and they are in fact governing themselves to the extent that the commands that they obey, whether based upon rules or not, are their own commands.” 16 The rules that govern Pauline’s behaviour are her own, and that is why she appears to be autonomous.

Furthermore, I would also argue that a person’s life-projects are normally linked to social roles and to the schemata that individuals associate with social roles. To be a painter, a farmer, an

16 Frankfurt, “Autonomy, Necessity and Love”, p. 131 It is only fair to note that Doris’ own vision of autonomy and responsibility is explicitly derived from Frankfurt (2002, ch. 7), and I expect that he would be quite friendly to these observations.
academic or a car mechanic is to identify with (and thus internalize) a general set of norms and values which provides the individual criteria for successful performance of that role. By acknowledging and identifying with such a schema, a person internalizes its norms in a very important way: the norms become part of the guiding principles that help to constitute her will.

Internalized role-schemata are thus deep features of agency as we know it. Furthermore, it is not conceivable that social roles could exist without collectively internalized clusters of behavioural norms. “Norms and norm-obeying behaviour,” writes Raimo Tuomela, “represent both a surface phenomenon of social life and a rock-bottom part of what the social order involves.”\(^{17}\) We therefore have good reason to think that both individual and social flourishing depend on the operation of role-schemata.

Certainly, there is much more to being an agent than simply adhering to internalized role-stereotypes. Nonetheless, I think significant progress has been made vis a vis the question we just asked: what general materials remain with which we might evaluate persons? My answer to this question is simple: we can evaluate persons as we always have, in light of the constitutive aims and norms associated with social roles.

**Social Roles and Evaluation**

Tuomela notes, correctly in my view, that internalized role-specific norms are necessarily connected to strong feelings of self-evaluation.\(^{18}\) A person’s successful adherence to the norms and aims of an internalized scheme produces feelings of self-esteem, whereas a failure to comply with role-specific rules or a failure to achieve role-specific aims causes feelings of shame. These observations suggest that social roles will help to ground what Williams called *internal* reasons

---

17 Tuomela 1995.  
18 Tuomela 1995, p.15.
for action. Furthermore, we can expect that in a given social context there will be a large amount of agreement on the basic norms and aims associated with various roles, and I would suggest that the process of socialization is, to a significant degree, the process of imparting this information to maturing persons.

In saying all of this, I am trying to show that the constitutive norms of social roles provide us with shared standards for ethical evaluation, standards that can have powerful motivational efficacy for individuals. While I do not expect that these standards, on their own, can deliver a fully comprehensive moral theory, it is important to keep in mind that this system of evaluation is one that is likely to be supported by situationist social psychology. Indeed, experiments like the one conducted by Liberman, Ross and Samuels provide us with reasons to think that role-based evaluation can be grounded in some very deep facts about human beings and the way they behave. We also have more traditional, ‘armchair’ reasons to think that role-based evaluation can work: we know that social roles exist, and such roles could not exist unless they were occupied by persons who usually conform to the relevant constitutive norms in role-relevant situations. So, we can expect that when persons are acting as representatives of a role with which they identify, they will be at least somewhat resistant to apparently trivial situational pressures.

Furthermore, role-based evaluation can be grounded in just the kind of explanatory and predictive power that local characterological discourse was meant to capture. To say that Pauline is a good mother is not just to provide an evaluation of her behaviour, it is also to explain that same behaviour in terms of her internalization of certain schema-based norms and values, and to predict that these norms and values will possess motivational authority for her in the future.

20 Though, of course there will never be total consensus, and much of the political/ethical discourse in a given society can be expected to revolve around just this sort of disagreement. Nonetheless, some basic agreement seems to be a prerequisite for the existence of societies as such.
Notice, however, that this prediction is implicitly situational, for we do not mean to say that she will be constantly moved to care for her children in all situations. Rather, we mean to say that the role-schema will motivate her in role-relevant situations. If situationist experiments showed that persons did not reliably act in service of the norms and values associated with their ground projects, then I would be forced to abandon this position. However, I am quite confident that such experiments cannot be produced, because I believe that the internalization of role-schemata is a necessary condition for the social order as we know it.

Whether there is some objective fact of the matter concerning these norms, or whether they are, as is sometimes said, “socially constructed”, is a question for future study. 21 Role-based evaluation is not purely relativistic, as there is significant logical space within which we might objectively criticize various schemata. The breakdown of traditionally restrictive gender-role schemata appears to be a result of just this sort of criticism. People once believed that women were just not capable of high intellectual, artistic or political achievement, and social interaction was deeply structured by the shared acceptance of this schema. However, this belief is, for obvious reasons, no longer credible, and those historical figures who doggedly insisted that it was false did not simply win the argument, they were objectively correct.

Conclusions

I think that reading Doris and the situationists can induce a kind of aporia, that vertiginous feeling that arises when we encounter a deep tension or contradiction in our thought. The contradiction, as I see it, is this: human life as we know it would not be possible without collective adherence to a large set of behavioural norms. We constantly depend on one another to

---

21 Ian Hacking’s The Social Construction of What? remains an invaluable guide to such questions, and I take his work to be directly relevant the view I am advocating in this paper.
exhibit this behavioural regularity, building projects and relationships on it, even occasionally staking our lives on it. Yet, to read the results of the situationist experiments is not just to be confronted with the thought that human behaviour is radically situation-dependent, but with the fact that the things which can exert commanding influence over human behaviour are shockingly trivial. I seem to be immersed in a world where I can trust my others (and myself) to adhere to certain modes of speech, feeling and behaviour, yet if our willingness to help others can be radically affected by the prior discovery of ten cents in a phone booth, then it is not clear how my confidence in myself and in others could possibly be justified. We know that this can’t be the whole story, yet there it is.

Philosophy is therapy, so the Wittgensteinean saying goes, and the analysis I’ve given in this paper is an attempt to therapeutically mollify this disorientation, a feeling which Solomon seems to have experienced quite acutely. I hope that this paper can both alleviate those symptoms and shed some light on how ethical thought can proceed in light of Doris’ powerful assault on globalism. As I tried to suggest, our trust in friends, spouses and colleagues is underwritten by the fact that they care about us, not by their possession of some very fine-grained disposition. In public life, the massive behavioural regularity we experience is real, and I suspect that it is largely a result of shared, internalized role-schemata. Renewed attention to the conceptual materials supplied by social roles should remind us that even with the elimination of global characterological discourse we still possess a huge number of contextual evaluative concepts. Indeed, they have been there all along. We can say, with confidence, that Pauline is a good mother and a bad co-worker and a so-so spouse, and all of these descriptions can contribute to a picture of what can still be called her character. Nonetheless, perhaps we ought to be suspicious of what might be called the globalist ambition: to discover a single, non-contextualized metric with which we can compare Pauline to other persons as such.