LUDWIG'S PUNCH AND BERTIE'S COMEBACK: RECONCILING RUSSELL AND WITTGENSTEIN ON THE CONTENT OF DESIRES

Peter Baumann
Philosophy / Swarthmore College
Swarthmore, PA 19081, USA
pbauman1@swarthmore.edu

Desires are contentful mental states. But what determines the content of a desire? Two different classic answers were proposed by Russell and by Wittgenstein, starting in the 1910s. Russell proposed a behaviourist account according to which the content of the desire is fixed by the type of state that puts an end to the relevant kind of behaviour which was triggered by some initial discomfort. The desire's content consists in its "satisfaction conditions". Wittgenstein criticized such an account for neglecting the crucial point that the relation between a desire and its content is a conceptual, internal one, not an external contingent one. Desires specify their own contents, their "fulfillment conditions". Even though there is a lot to say in favour of Wittgenstein's criticism, this paper argues that Russell pointed at an important aspect of desires which plays a crucial role for accounts of self-knowledge of one's own desires. It turns out fulfillment conditions and satisfaction conditions are tied together in rational self-knowledge of one's own desires. In this sense, the views of Russell and Wittgenstein can be combined in a fruitful way.

1. DESIRES: TWO VIEWS

Desires are intentional mental states: they have content. The content of a desire is part of what makes a desire the state it is; the other part is played by the fact that it is a desire and not some other type of state, like, e.g., a belief or a fear. But what is the content of a desire and what fixes it? What kind of relation is there between a desire and its content? Similar questions can be raised

n.s. 40 (winter 2020): 132–49
ISSN 0036–01631; online 1913–8032

about preferences and related "pro-attitudes"; for the sake of simplicity we can restrict ourselves to desires here.

A first answer goes like this. Desires are propositional attitudes. It is true that we often talk as if desires are directed towards objects ("I want a drink"), activities ("I want to swim a bit now") or events ("I want this to end"). However, I am assuming here, without further discussion, that all these ways of talking turn out to be propositional in nature if further specified ("You want a drink? What for? Is it that you want to quench your thirst with it? Pour it over my head? Oh, I see: you want that the barkeeper is happy!"). The content of a propositional attitude is the content of the relevant proposition or—if you do not believe in propositions—what the that-clause of the corresponding sentence describing the attitude specifies. The content of Annie's want for a ham sandwich is that she gets a ham sandwich. The content of a desire determines what "fulfills" it: "fulfilling a desire" means making its propositional content true. Thus, the first answer to our question can be put like this: the content of a desire is its fulfillment condition (its content's truth condition).[1]

Strictly speaking, a particular desire, like Ann's desire that she have a ham sandwich, has a set of many fulfillment conditions (having a ham sandwich in front her, or having it while the radio is on, or while the sun is out, etc.). For simplicity's sake we can speak simply of a desire's fulfillment condition as the set of all those more specific conditions. It should be added here that many reports of desires seem to underspecify the content: even though it might be true that Mary wants the quarrel between her friends to end, it might not be true that Mary wants the quarrel between her friends to end by shooting them.[2] Whether it is the semantics or the pragmatics of desire ascriptions that should deal with this problem and how it should be solved can be left open here (perhaps some of the content need not even be represented by the subject[3]).

There is a fundamentally different answer to our question about desires and their contents. According to this second view, propositional

[1] There is no commonly accepted and shared use of terms like "fulfillment" and "satisfaction". The explanations given here should prevent any misunderstanding.

[2] See for this kind of issue Fara, "Specifying Desires" (2013), as well as Shier, "Direct Reference for the Narrow Minded" (1996), p. 227 *passim*, who distinguishes between characterizing and specifying the content of a desire.

[3] See Perry, "Thought without Representation" (1986).

content does not matter much or at all here. Desires are rather seen as certain states of the person or the organism and as typically (though not always) having, *inter alia*, certain behavioural effects. To give an example, Fred feels cold and approaches the heater; he soon starts to feel warm. The content of a desire, then, is identified with that state of affairs which puts an end to the state that caused the behaviour. It is typically (though not always) the cause of a state of satisfaction in the person or organism. It constitutes the desire's satisfaction condition rather than its fulfillment condition (we can put cases aside here where there fails to be satisfaction; but see the next paragraph). Since there are many conditions that satisfy a desire we should rather say that the content is the set of all those conditions that satisfy the desire. For brevity's sake I will simply talk about a "satisfaction condition" here.

The fundamental difference between these two views about the content of desires[4] can be easily overlooked because often fulfillment goes hand in hand with satisfaction. Annie is hungry and she loves ham sandwiches: hence, getting a ham sandwich will as much fulfill her desire as satisfy it. However, fulfillment conditions and satisfaction conditions can diverge. What fulfills a desire need not satisfy it: Annie, after having eaten the ham sandwich she wanted to eat, might still be hungry or even more hungry than before, and she might even regret having gone for that ham sandwich. And, vice versa, what satisfies a desire need not fulfill it: even though Annie had a cheese sandwich instead of a ham sandwich it satisfied her wishes even better than a ham sandwich could have (which might not have satisfied her at all).

The satisfaction view of content emphasizes the motivational aspect of desires whereas the fulfillment view rather focuses on the intentional ("aboutness") aspect of desires. Some people, like Jerry Fodor for instance, thought that it is one of the main tasks in the philosophy of mind to bring these two aspects together and to explain how intentional states can motivate behaviour.[5] I will not go into this issue here.

According to the satisfaction view, desires are pretty close to drives and at least some of them are drives. Non-human animals certainly have desires in that sense—and humans probably, too, to some degree.

[4] See also, more generally MÁCHA, *Wittgenstein on Internal and External Relations* (2015), Ch. 11.

[5] See, e.g., FODOR, "Propositional Attitudes" (1981), pp. 182ff.; see also DRETSKE, *Explaining Behavior* (1988), p. 130, for the difference between fulfillment and satisfaction.

Such desires need not be mental states in anything more than a primitive or rudimentary sense. According to the fulfillment view, desires are more sophisticated mental states, perhaps even requiring the mastery of a natural language.

According to the fulfillment view but not according to the satisfaction view, attributions of desires are referentially opaque. According to the former view, it might be true that Stephane wants a piece of the cake on the left but it might not be true that Stephane wants a piece of the last Linzer Torte the chef produced last night (given that the cake on the left is yesterday's last Linzer). To be sure, this holds for *de dicto* ascriptions of desires but not for *de re* ascriptions: if it is true that it is from the cake on the left that Stephane wants to eat, then it is also true (as we may assume for the sake of the example) that it is from the last Linzer Torte which the chef produced last night that Stephanie wants to eat—and vice versa. However, we can put this aside here because only *de dicto* ascriptions capture the kind of content that we're interested in here. To say that Chuck desired of the hemlock to drink it, doesn't specify what he wanted: to commit suicide or to quench his thirst, or something else altogether.

Also, according to the fulfillment view, desires support inferential relations with other attitudes, especially other desires and beliefs. This allows for an account of practical reasoning and rationality. Suppose I want to get into city centre quickly. Suppose further that I believe (correctly, let us assume) that the best way for me to get into city centre quickly is to take a train. As a rational person, I should then also want to take a train. My practical syllogism that I should take a train because doing this is the best way to get me where I want to be is a good one. The satisfaction view on the content of desires does not seem to allow for, or at least not suggest, this type of account of practical reasoning; it is certainly not easy to see how it could allow for or support any such account.

One might be tempted to think that there are two different kinds of desires at issue here. However, I do find this implausible. There is just one type of mental state called "desire" (in all its variations, of course). The relevant question rather is: What is the right view about the content of desires? The fulfillment view or the satisfaction view or a combination of both or none of these?

2. RUSSELL

Bertrand Russell—at least at some points in his long career—thought that the satisfaction view gives the right answer to the questions above about the content of a desire. Ludwig Wittgenstein famously criticized a general account of propositional attitudes as well as a view on desires like the ones Russell defended. In opposition to Russell, Wittgenstein argued for the fulfillment view. I shall focus here on the disagreement between Russell and Wittgenstein.[6] This is not merely of historical but also of systematical importance. It is interesting to everybody who is involved in the current debates on content. Let us consider Russell's view first.

Starting in the 1910s and then in particular in the preliminary studies to *The Analysis of Mind* around 1919,[7] in *The Analysis of Mind* (1921), and in *An Outline of Philosophy* (published 1927) Russell puts forward a strongly behaviourist theory of motivation and action.[8] According to this theory, desires are to be explained via behaviour cycles: A particular bodily state of an organism (e.g., lack of nutrition) expresses itself in feelings of discomfort (e.g., hunger) and triggers a certain type of behaviour (e.g., approaching "foody" regions). This behaviour is "directed towards" or better, brings about a state of affairs that stops the behaviour by ending its triggering conditions, eliminates discomfort (e.g., by getting food) and brings about pleasure. Thus,

[6] One should keep in mind that as far as available texts are concerned, Wittgenstein's critique started before Russell began to defend the kind of view Wittgenstein criticized. I will not go here into questions about what has or might have been said in some of the many conversations between Russell and Wittgenstein in the 1910s. When I talk about "Wittgenstein's critique of Russell" in the following, I don't want to imply a certain temporal order but rather intend to refer (for lack of a better and equally concise expression) to Wittgenstein's critique of the kind of view Russell defended for some time.—It is an interesting and important historical question whether the different views of Russell and Wittgenstein resulted, roughly, from their different responses to the problems of the multiple-relation theory of judgment that Russell held earlier (thanks to a referee for making this point!). I am not going to go into this here because this would lead to far away from the main aim of this paper; the focus here is not on the prehistory of the different views Russell and Wittgenstein held about the content and nature of desires.

[7] See RUSSELL, "The Anatomy of Desire" (pts. I–III, 1919), as well as "Analysis of Mind" (1919) and "Miscellaneous Notes" (1919), published as **1** and **2** in *Papers* 9.

[8] See RUSSELL, *An Outline of Philosophy* (1927), pp. 32ff., 93ff., 229ff.; *AMi*, pp. 59ff., 30off., and "The Anatomy of Desire" (which largely overlaps with *Ami*); "Analysis of Mind", pp. 11ff., "Miscellaneous Notes", p. 17ff.

according to Russell, the content of a desire consists in the "halting" conditions of the behaviour triggered by the desire. In other words, the content of a desire is its satisfaction condition. We can see here already that Russell's account is not completely behaviourist since he allows for feelings, images and other "inner" states. But the behaviourist element in his view is very strong.[9]

Russell says: "A hungry animal is restless until it finds food; then it becomes quiescent. The thing which will bring a restless condition to an end is said to be what is desired."[10] And, in more detail:

> A mental occurrence of any kind … may be a cause of a series of actions, continuing, unless interrupted, until some more or less definite state of affairs is realized. Such a series of actions we call a "behaviour-cycle"…. The property of causing such a cycle of occurrences is called "discomfort"; the property of the mental occurrences in which the cycle ends is called "pleasure"…. The cycle ends in a condition of quiescence, or of such action as tends only to preserve the *status quo*. The state of affairs in which this condition of quiescence is achieved is called the "purpose" of the cycle, and the initial mental occurrence involving discomfort is called a "desire" for the state of affairs that brings quiescence.[11]

Russell does thus not individuate the content of a desire by indicating its propositional content but rather by referring to its satisfaction condition.

Since it is a contingent fact that conditions of type *A* satisfy organisms of type *B* if they are in states of type *C*, we can also call Russell's account an "externalist" one. It also has certain similarities with causal functionalism in the philosophy of mind:[12] According to the latter view, desires are states individuated by characteristic types of causes, effects and relations to other mental states; according to Russell, desires are states individuated by triggering conditions causing some relevant type of behaviour and by halting conditions causing the end of that behaviour. Even though Russell only mentions the latter conditions in the quotation above, we need to add the former given that the

[9] See also Griffin, "Russell's Neutral Monist Theory of Desire" (2015), p. 10, or Kenny, *Action, Emotion and Will* (2003), p. 73.
[10] *AMi*, p. 32; see also "Analysis of Mind", p. 18.
[11] *AMi*, p. 75; see also p. 63f.; "Analysis of Mind", pp. 13f., 11ff.; *OP*, p. 93.
[12] See, e.g., Block, "Troubles with Functionalism" (1980).

same conditions can be halting conditions for different types of behaviour triggered by different kinds of states: Finding refuge inside a bank satisfies those who try to escape particularly bad weather as well as those who are reluctant to be seen by a nosy and chatty neighbour they just spotted on the street.

It should also be added briefly here that a Russellian account has to be careful to distinguish in principle between behaviour cycles and other types of feedback mechanisms (like a thermostat regulating the temperature in a room): Do the feelings of discomfort and pleasure make all the difference here? This points to a problem for Russell's account. He discusses it briefly[13] and invokes the availability of physical or mechanical explanations in the case of non-behaviour (natural processes or involuntary movements)—which makes one wonder why such explanations are not available in the case of the explanation of behaviour. We can leave the question open here whether this problem admits of a solution within the confines of Russell's view.[14]

One further issue deserves to be mentioned at least briefly here. One might wonder what exactly counts as a desire, according to Russell. Is it the discomfort that triggers the whole cycle? This would identify desire with a certain kind of feeling and not require any reference to behaviour; furthermore, feelings are not fine-grained enough to capture all the subtle differences between desires. Should one then identify desires with the pleasure at the end? This seems implausible on the face of it. Again, it would be hard to see how pleasure could be so fine-grained as to capture the difference between desires: the desire that *A* win the election in country *X* is different from the desire that *B* lose the election in country *X*, but the pleasure might be just one and the same. Is the desire then to be identified with the whole cycle? Perhaps it is sufficient here to attribute to Russell the view that talk about a particular desire has the same truth conditions as talk about a particular behaviour-cycle. The further details can be left open here because we are rather interest in the notion of a content of some desire.

Russell's view has some interesting implications which I want to mention briefly. According to Russell, we know our own desires in the same way in which we know other people's desires: by observing behaviour and its cycles and by making inferences to the desire. There is

[13] See *AMi*, pp. 66–7.
[14] See also KENNY, pp. 74–5.

no "special" kind of experience available exclusively from the first-person perspective (as opposed to the third-person perspective).[15] There is no immediate or incorrigible self-knowledge here.[16] Russell says: "I believe that the discovery of our own motives can only be made by the same process by which we discover other people's, namely, the process of observing our actions and inferring the desire which could prompt them."[17] And, with regard to the satisfaction of a desire he says: "But only experience can show what will have this sedative effect, and it is easy to make mistakes."[18] Thus, Russell's example suggests that a satisfaction view goes hand in hand not only with an anti-Cartesian but also with an anti-commonsensical stance on self-knowledge (assuming that common sense has it that we have special access to our own mental states).

3. WITTGENSTEIN'S PUNCH AND RUSSELL'S ISSUES

I do not intend to go into the historical details of Russell's many debates with Wittgenstein and describe how any real debate between them went. Rather, I am interested in Wittgenstein's critique of a Russellian conception of desire. It started in the *Notebooks* from 1914–16.[19] As is well known, Wittgenstein seems to have a "knock-down" argument against the kind of view Russell defended: "If I wanted to eat an apple, and someone punched me in the stomach, taking away my appetite, then it was this punch that I originally wanted."[20] This objection is supposed to show that the conditions of cessation of a desire cannot determine the content of the desire.

According to Wittgenstein, the relation between a desire and its content is not an external, empirical one (as Russell thought) but an in-

[15] See *An Outline of Philosophy*, p. 229f., and *AMi*, p. 31f.
[16] See *AMi*, pp. 30ff.
[17] *AMi*, p. 31.
[18] *Ibid.*, p. 32.
[19] See *Notebooks*, p. 77; *Philosophical Remarks*, §§22ff.; *Wittgenstein's Lectures: Cambridge 1930–32*, pp. 9, 30, 33; *Philosophische Grammatik*, pp. 134, 138, 140, 150f., 157f.; *Blue and Brown Books*, pp. 21f., 37f.; *Philosophical Investigations*, §§440f., 451ff., 458ff.; *Zettel*, §§53, 68.
[20] *Philosophical Investigations*, §22; see *Wittgenstein's Lectures: Cambridge 1930–32*, p. 33; see also GRIFFIN, p. 15, on the case of failure: wouldn't Russell have to say, very implausibly, that the unintended and unwanted consequence is what was desired?

ternal, logical relation. Wittgenstein says: "Saying 'I should like an apple' does not mean: I believe an apple will quell my feeling of nonsatisfaction. This proposition is not an expression of a wish but of nonsatisfaction."[21] He says similar things in the *Blue Book*:

> It is a hypothesis that a particular feeling of hunger will be relieved by eating a particular thing. In Russell's way of using the word "wishing" it makes no sense to say 'I wished for an apple but a pear has satisfied me'. But we do sometimes say this, using the word "wishing" in a way different from Russell's. In this sense we can say that the tension of wishing was relieved without the wish being fulfilled; and also that the wish was fulfilled without the tension being relieved. That is, I may, in this sense, become satisfied without my wish having been satisfied.[22]

The relation between a desire and its content is an internal, logical one: "A wish seems already to know what will or would satisfy it; a proposition, a thought, what makes it true—even when that thing is not there at all!"[23] And in the Cambridge Lectures from 1930–32, Wittgenstein says: "What you now wish or expect is not a matter of future experience; for your wish or expectation may never be fulfilled."[24] Thus, Wittgenstein holds an "internalist" conception of desires according to which the content of a desire consists in the propositional content of this practical attitude, i.e., in its fulfillment conditions.

Wittgenstein, too, draws some conclusions as to the nature of self-knowledge. But different writings suggest different ones. In the *Blue Book* Wittgenstein characterizes utterances like "I know what I want" as grammatical utterances[25] In the *Investigations*, however, he claims: "'I know what I want, wish, believe, feel, …' (and so on through all the psychological verbs) is either philosopher's nonsense, or at any rate not a judgment a priori."[26] The fulfillment view as such does not seem

[21] *Philosophical Investigations*, §440.

[22] *Blue and Brown Books*, p. 22; see *Philosophical Investigations*, §441.

[23] *Philosophical Investigations*, §437.

[24] *Wittgenstein's Lectures: Cambridge 1930–32*, p. 33; see *Blue and Brown Books*, p. 37f.

[25] See *Blue and Brown Books*, p. 30.

[26] *Philosophical Investigations*, p. 221; see also here: McGuinness, "I Know What I Want" (1956–57), pp. 305ff., and von Savigny, "I Don't Know What I Want" (1982), pp. 193ff.; concerning the question whether a person usually or always knows what she wants and whether there is a difference between the first-person perspective and

to imply any particular view about first person access to the content of desires.

Is Wittgenstein right in his critique of Russell's theory of desires? Wittgenstein's remarks like the one about the wish for an apple and the punch in the stomach show at least two different things. First, Russell's conception of satisfaction is much too broad and not satisfactory. A punch in the stomach may make my wish for an apple cease but it does not satisfy my wish—or at least, it very often does not do that. The cessation of a desire is something different from both the satisfaction and the fulfillment of a desire. If I, for instance, finally resign and give up my wish to get some Spaghetti Carbonara today, my desire might peter out, but it is, of course, neither fulfilled nor satisfied. Russell's theory of desire thus needs to be supplemented with at least a threefold distinction between cessation, satisfaction, and fulfillment of a desire. So, what is the difference between cessation and satisfaction?

I do not see why one should not, in principle, be able to repair this part of Russell's theory. Satisfaction has an emotional or affective side that mere cessation (without satisfaction) lacks: the organism is pleased or relieved of discomfort, and this is an effect of the behaviour initiated by discomfort. Satisfaction is a special case of cessation. Perhaps Russell's "pleasure-condition" that the cessation of the relevant behaviour is pleasant to the organism (see above) can already take care of this point.

This also takes care of the worry that unrelated events might lead to pleasure or relief of discomfort. For instance, unbeknownst to me and not triggered by me, someone might have put a happy-making substance in my coffee. I drink my coffee and am very happy as a result. There are many functional equivalents to the agent's behaviour which are not part of the their behaviour. If I don't scratch my itchy elbow, someone else might. Even some event that does not constitute behaviour at all might do the job (a branch of a tree falling on my itchy elbow, taking away the itch). All this can happen, but in those cases the causal genealogy is not the relevant one: the pleasure or the relief of discomfort is not caused by some agent's behaviour which itself was caused by the agent's discomfort. It is a different story, of course, if my (intentional) making unhappy faces makes my close companion

the third-person perspective here, see: *Blue and Brown Books*, p. 22; *Philosophical Investigations*, §§441, 587; *Wittgenstein's Lectures: Cambridge 1930–32*, p. 183.

put some of their happy-makers into my coffee.

But there is another part of Russell's account—a part that seems beyond possible repair. This leads to Wittgenstein's second, main point against Russell: the relation between a desire and its content is an internal, logical one and not just an external, contingent, and "merely" causal relation. What Wittgenstein's objection so convincingly shows (even if Wittgenstein does not spell it out in detail and does not offer an alternative account) is that Russell leaves out something important: he over-emphasizes the motivational aspect of desires and neglects their propositional attitude-aspect.[27]

Closely related to this problem is the fact that talk about desires creates intensional contexts whereas talk about states of an organism and its behaviour doesn't. I might show theatre-visiting behaviour, but what is it about theatre performances that attracts me (pulls or pushes me there)? Is it the staging of some drama? Or rather the literary qualities of certain plays? A behaviourist theory will tend to be way too coarse-grained in its analysis of desire in order to capture the fine-grained, intensional structure of desire.[28]

There are more issues for Russell's account (not mentioned by Wittgenstein) which bring up more of such serious problems. It seems that not all desires lead to action or behaviour.[29] My desire to travel to Paris (from far away) within ten minutes and without much cost does (due to my knowledge of the relevant facts) not lead to any specific behaviour. I might, of course, show linguistic behaviour and utter sentences like "Oh, how nice that would be!" but then again, this need not be the case. It also does not help to insist that omissions as well as acts must be counted as behaviour: What action do I not perform because of the above desire? These problems arise for desires about the past or desires the fulfillments of which are impossible in some sense. A related problem arises in cases of conflicts of desires: both when the subject ends up paralysed and when one desire wins over the other.

One might respond that perhaps one can identify relevant counterfactuals and dispositions to behaviour which are characteristic of a given desire, no matter whether the disposition is ever manifested or

[27] Griffin, pp. 13–14 discusses a statistical interpretation of Russell's view (the discomfort only "tends" to trigger behaviour of a certain sort) as a reply to Wittgenstein's "punch" but rightly finds this wanting, too.

[28] See also Kenny, pp. 75–6, on this as well as Griffin, p. 22.

[29] See, e.g., Bostock, *Russell's Logical* Atomism (2012), p. 187.

not. My desire to travel to Paris in 10 minutes and without much cost would make me smile and say "yes!" if I encountered a credible offer (which I never will). One drawback of this proposal, however, is that it seems to change the theory very much and turn it into an altogether different theory. A desire does not trigger a disposition to behave a certain way but rather the behaviour itself. According to this proposal, desires would rather be taken to consist, partly, in such dispositions to behave a certain way. It is worth adding that Russell's naturalism would have made him at least uncomfortable with such a counterfactualist extension of his view.[30] Apart from this, it is not clear how this could work in the case of desires for impossible things. It might be that the best way to rescue Russell's account from this kind of problem is to restrict it to desires with behavioural effects. But this loss of generality is a high price to pay.

There are also serious circularity worries about Russell's account. A desire for *X* is individuated and identifiable at least in part via a certain kind of behaviour, one that leads to *X*. But can one in turn understand the behaviour that leads to *X* in a "bare bones" behaviourist way, that is, without understanding it as intentional, as behaviour motivated by a desire for *X*? Consider this case.[31] Abe wants to kill Kay while climbing. This makes Abe so nervous that he loses the grip on the rope which makes Kay fall to his death. Contrast this with a case where Abe loses the grip "intentionally" (this relates to be point above about unrelated functional equivalents). Only in the second case do we have a piece of behaviour relevant to the satisfaction of the desire. Can one distinguish behaviour from non-behaviour or relevant behaviour from irrelevant behaviour in a strictly behaviourist way without using the "intentional idiom"? Some might try to explain the difference using a notion of deviant causation; however, so far convincing accounts have been missing. I don't want to rule out here that there can be solutions to this circularity worry but rather point to serious doubts one should have.

Harder to solve and closely related to this is another problem. Russell gives the following definition of "behaviour-cycle": "A 'behaviour-cycle' is a series of voluntary or reflex movements of an animal, tending to cause a certain result, and continuing until that result is cause,

[30] See Kitchener, "Bertrand Russell's Flirtation with Behaviorism" (2004).
[31] See Davidson, "Freedom to Act" (1980), p. 79.

unless they are interrupted by death, accident, or some new behaviour-cycle."[32] There are certainly criteria for death that are independent from behaviourial criteria but what about accidents? What constitutes an accident depends on what constitutes a behaviour-cycle; so, the latter cannot be individuated in a non-circular way via the former.[33] Similar problems arise for the individuation of behaviour-cycles. Nothing in the behaviour itself seems to determine whether there is one longer cycle or rather an interrupted cycle followed by a complete cycle.[34]

One might reply that this type of circularity is not vicious and point out that many concepts are mutually explanatory in a non-problematic way. However, the circularity we find here concerns terms of a theory which are at least to a great degree technical terms. Hence, the circularities above stand in the way of clarifying and understanding these terms as well as making use of them in illuminating explanations. In this respect it is a bad circularity. Therefore, Russell cannot just declare Wittgenstein's punches to be mere accidents.

One may also wonder what makes the causal connection between discomfort, behaviour and resulting pleasure. In the case of unlearned behaviour this would just be "wired in". But so much human or animal behaviour is learned. It is hard to imagine how the connection could be made without some learned representation of one's environment,—some beliefs the organism has acquired and which it applies in specific situations. No learned behaviour seems conceivable without this kind of cognition. Adding this to the theory makes it much less behaviourist and much more "cognitivist". It is also already hard enough to see how discomfort or pleasure could be explained behaviouristically and in a non-circular way.[35]

Finally, there are some limitations which have already been mentioned above. Russell's view cannot take account of the fact that standard, *de dicto* attributions of desires are referentially opaque: "Annie wants a ham sandwich from the store next door" does not imply "Annie wants one of the cheapest sort of sandwiches sold in the store next door", even given that ham sandwiches in fact are the cheapest ones

[32] *AMi*, p. 68.
[33] See GRIFFIN, pp. 16–21.
[34] See KENNY, p. 75, and GRIFFIN, pp. 21–5.
[35] See also KENNY, pp. 73–4.

you can get next door. Furthermore, Russell's view cannot explain the inferential relations of desires with other states: For instance, it cannot explain how (i) wanting to get a ham sandwich and (ii) believing that they sell them in the store next door can lead rational Annie (iii) to want to go over to the store or even to go over to the store. How could we ever get to an account of practical reasoning on the basis of Russell's view of desires?

It seems then that we have to choose between two ways of looking at Russell's account of desires. We can see it either as an at best incomplete but probably untenable view of desires and their contents in general. Or we can see it as a more complete view of one very special type of desires, namely basic drives. Since the latter option was not Russell's aim and is not very interesting due to its restriction, we have good reasons to favour the first interpretation. This, however, must appear very unwelcome to any Russellian.

4. PUTTING RUSSELL AND WITTGENSTEIN BACK TOGETHER AGAIN[36]

Does all this mean that Wittgenstein "wins" and Russell "loses" the philosophical ping-pong game? Do satisfaction conditions have nothing at all to do with the content of desires? No, it is not that simple. To be sure, the relation between a desire together with its content (respecting fulfillment) on the one hand and its satisfaction on the other hand is an external, contingent one. But isn't there still a close, internal connection between (a) having a belief about one's desires and their contents and (b) having a belief about the satisfaction conditions of one's desires? If not, the following sentence—taken as an expression of a belief about one's desires conjoined with a belief about that desire's satisfaction conditions—would have to be acceptable: "I want to eat an apple but that won't satisfy me!" However, such an utterance sounds quite infelicitous or even "paradoxical". One might even be reminded of Moore-paradoxical propositions.[37]

It is remarkable that in the *Remarks on the Philosophy of Psychology*

[36] The aim of this section is not to develop a more Wittgenstein-friendly Russellian view in any detail (e.g., in the light of the above-mentioned non-behaviourist elements in Russell's thinking at the time). My aim is rather to argue for the possibility of a "reconciliation" between the two views.

[37] See MOORE, "Russell's Theory of Descriptions" (1944).

Wittgenstein himself considered a similar sentence.[38] One can, of course, say something like "I want an apple but that won't satisfy me completely because I could eat much more than that!" or "If I'm not allowed to have a steak, and if you force me to choose between fruits, then I want an apple even though that won't satisfy me at all and only dissatisfies me a bit less than the other fruits!" Cases like these are, of course, unproblematic. However, what the "paradoxical" case above shows is that it would be irrational not to give up a wish if one is convinced that its fulfillment would not be satisfying at all or choice-worthy in any way. "I want to eat an apple but that won't satisfy me!" is paradoxical if interpreted in the sense of "I want to eat an apple but that won't satisfy me in any way!" or of "I want to eat an apple but that will dissatisfy me more than anything else on the menu!" Rational persons cannot believe they have a certain desire without also believing that its fulfillment will be satisfying (in the relevant sense). One could perhaps even go one step further and claim—based on a theory of radical interpretation like the one defended by Donald Davidson[39]—that the case of a subject believing both conjuncts (the one about the desire and the one about the lack of satisfaction) is not even a case of irrationality but rather a case of "collapse of belief". Such a person could not be taken to have beliefs about a given desire, its fulfillment and its satisfaction, in the first place. However, I don't want to make such strong assumptions about interpretability here and will rather go with the idea that there is a rationality constraint binding beliefs about desires and beliefs about satisfaction together.

Consider the example of a reluctant drug addict who wants to stop taking the drug but simply cannot stop taking it.[40] Even if he believes

[38] See *Remarks on the Philosophy of Psychology*, I: 492: "Denk dir, Einer sagte "Ich wünsche,—will aber nicht, daß mein Wunsch befriedigt werde.—(Lessing: "Wenn Gott in seiner Rechten …") Kann man also Gott bitten, den Wunsch zu geben, und ihn *nicht* zu erfüllen?" (my translation: "Think of someone who said "I want,—but don't want, that my wish will be met.—(Lessing: "If God in his right hand …") Can one thus ask God, to give us the wish but *not* to meet it?"). The sentence Wittgenstein mentions here is not quite of the kind discussed above but similar. The German verb "*befriedigen*" when applied to a wish might sound as if it refers to satisfaction but it doesn't: it refers to the fulfillment ("*erfüllen*") of a wish. To self-ascribe a wish while also wishing for its non-fulfillment might be possible only in special cases like ones of addiction (see below).

[39] See DAVIDSON, "Incoherence and Irrationality" (1985).

[40] See, amongst others, FRANKFURT, "Freedom of the Will and the Concept of a Person" (1988), p. 21.

that the fulfillment of his addictive desires is satisfying in some respect, this belief is "cancelled out" by the belief that, all things considered, the fulfillment of his addictive desires is not satisfying (more dissatisfying than satisfying). Hence, he is irrational not only because he desires something that, according to his own best judgment, will not satisfy him. He also believes that he desires something that, according to his own best judgment, will not satisfy him. All this illustrates that for rational people the belief that one has a desire that *p*, comes with the belief that *p* would be satisfying for oneself. Desire and satisfaction are, so to speak, internally connected via the belief rational people hold about the fulfillment and the satisfaction aspect of their desires; more precisely, the belief that one has a desire *D* and the belief that fulfilling *D* will be satisfying to some degree (and not too dissatisfying in other respects) are internally connected with each other—even though, as Wittgenstein rightly states, the relation between desire and satisfaction is a merely external one.

As time goes by, people learn what kinds of states of affairs are satisfying (and to what degree) and which ones not.[41] Among the things people learn is to produce behaviour with satisfying effects (Russell could argue like that[42]). Hence, by and by a person can come to wish that *p* because and insofar as she thinks that *p* being the case would be satisfying. Often desires are based on beliefs as to whether the fulfillment of the desire would be satisfying. The satisfaction aspect can even be part of the propositional content of the desire. It would be incorrect to assume that this is excluded by the concept of a desire's content. Hungry Annie might want to eat a ham sandwich "under the description"[43] that eating kills hunger, eliminates the desire to eat, and is satisfying. If the propositional content of a desire refers to the conditions of satisfaction of the desire, then the desire becomes partly self-referential. Then, the desire that *p* is at least partly a desire that itself be satisfied. This self-reflectivity, however, is not a universal trait of desires.[44]

[41] See Sobel, "Full Information Accounts of Well-Being" (1994), p. 788.

[42] See Russell, "Analysis of Mind" (1919), p. 12.

[43] See Anscombe, *Intention* (1958), sec. 23, pp. 46–8 *passim*.

[44] See Schiffer, "A Paradox of Desire" (1976), p. 199; about self-reflective intentions see Searle, *Intentionality* (1983), pp. 85ff., and Harman, "Practical Reasoning" (1976), p. 441.

5. CONCLUSION

Let me recapitulate. First, Wittgenstein rightly objects against Russell that the content of a desire is its propositional content and not its satisfaction condition. The internalist account of content according to which the content of a desire is not merely contingently related to the desire is the right one (as I said above: this does not rule out that conditions of satisfaction can be part of the propositional content). But second, Russell's view suggests or at least helps us develop an important point which Wittgenstein missed: desire, fulfillment and satisfaction are indirectly related to each other, namely via the beliefs of a rational person concerning them; for rational subjects, the belief that one desires that *p* comes with the belief that *p* would be satisfying. We can thus say that both, Russell as well as Wittgenstein, have not only made important points in the debate but could have reached a common view beyond dispute. Moreover, this common view constitutes an important insight into the rationality conditions pertaining to thoughts and beliefs about one's own desires,—an insight that deserves to be explored much more than it has been so far. The proposal above is a first step into that direction.

WORKS CITED

Anscombe, G. E. M. *Intention*. Oxford: Blackwell, 1958.

Block, Ned. "Troubles with Functionalism". In Block, ed. *Readings in the Philosophy of Psychology*, Vol. 1. Cambridge, MA: Harvard U. P., 1980.

Bostock, David. *Russell's Logical Atomism*. Oxford: Oxford U. P., 2012.

Davidson, Donald. "Freedom to Act". In his *Essays on Actions and Events*. Oxford: Clarendon P., 1980.

—. "Incoherence and Irrationality". *Dialectica* 39 (1985): 345–54.

Dretske, Fred. *Explaining Behavior: Reasons in a World of Causes*. Cambridge, MA: MIT P., 1988.

Fara, Delia Graff. "Specifying Desires". *Noûs* 47 (2013): 250–72.

Fodor, Jerry A. "Propositional Attitudes". In his *Representations: Philosophical Essays on the Foundations of Cognitive Science*. Cambridge, MA: MIT P., 1981.

Frankfurt, Harry G. "Freedom of the Will and the Concept of a Person". In Frankfurt, *The Importance of What We Care About: Philosophical Essays*, Cambridge: Cambridge U. P., 1988.

Griffin, Nicholas. "Russell's Neutral Monist Theory of Desire". *Russell* 35 (2015): 5–28.

Harman, Gilbert. "Practical Reasoning". *Rev. of Metaphysics* 29 (1976): 431–63.

Kenny, Anthony. *Action, Emotion and Will*. 2nd ed. London and New York: Routledge, 2003.

Kitchener, Richard F. "Bertrand Russell's Flirtation with Behaviorism".

Behavior and Philosophy 32 (2004): 273–91.

Mácha, Jakub. *Wittgenstein on Internal and External Relations: Tracing All the Connections*. London and New York: Bloomsbury, 2015.

McGuinness, Brian F. "'I Know What I Want'". *Proceedings of the Aristotelian Society* 57 (1956–57): 305–20.

Moore, G. E. "Russell's Theory of Descriptions" (1944). In his *Philosophical Papers*. London and New York: Allen & Unwin/Macmillan, 1959.

Perry, John. "Thought without Representation". *Proceedings of the Aristotelian Society, Supplementary Volume* 60 (1986): 137–51.

Russell, Bertrand. "The Anatomy of Desire: I. Desire in Animals". *The Athenaeum*, no. 4,676 (12 Dec. 1919): 1,340–1; **23** in *Papers* 15.

—. "The Anatomy of Desire: II. Primitive Desire in Human Beings". *The Athenaeum*, no. 4,677 (19 Dec. 1919): 1,372–3; **23** in *Papers* 15.

—. "The Anatomy of Desire: III. Conscious and Unconscious Desire". *The Athenaeum*, no. 4,678 (26 Dec. 1919): 1,402–3; **23** in *Papers* 15.

—. "Analysis of Mind" (1919); **1** in *Papers* 9.

—. "Miscellaneous Notes" (1919); **2** in *Papers* 9.

—. *AMi.*

—. *An Outline of Philosophy*. London: Allen & Unwin, 1927; repr. New York: Meridian Books, 1960.

—. *Papers* 9, 15.

Schiffer, Stephen. "A Paradox of Desire". *American Philosophical Quarterly* 13 (1976): 195–203.

Searle, John. *Intentionality: an Essay in the Philosophy of Mind*. Cambridge: Cambridge U. P., 1983.

Shier, David. "Direct Reference for the Narrow Minded". *Pacific Philosophical Quarterly* 77 (1996): 225–48.

Sobel, David. "Full Information Accounts of Well-Being". *Ethics* 104 (1994): 784–810.

von Savigny, Eike. "I Don't Know What I Want". *Grazer Philosophische Studien* 42 (1982): 193–209.

Wittgenstein, Ludwig. *Tagebücher/ Notebooks 1914–1916*. 2nd ed. G. H. von Wright and G. E. M. Anscombe, eds. Oxford: Blackwell, 1979.

—. *Philosophical Remarks.* Rush Rhees, ed. Trans. Raymond Hargreaves and Roger White. Oxford: Blackwell, 1964.

—. *Wittgenstein's Lectures*: *Cambridge 1930–1932*. Desmond Lee, ed. Totowa, NJ: Rowman and Littlefield, 1980.

—. *Philosophische Grammatik*. Rush Rhees, ed. Frankfurt am Main: Suhrkamp, 1973.

—. *Preliminary Studies for the "Philosophical Investigations" Generally Known as the Blue and Brown Books*. Oxford: Blackwell, 1958.

—. *Philosophical Investigations.* G. E. M. Anscombe and Rush Rhees, eds. Trans. Anscombe). Oxford: Blackwell, 1953.

—. *Remarks on the Philosophy of Psychology*, Vol. 1. G. E. M. Anscombe and G. H. von Wright, eds. Trans. Anscombe. Vol. 2. von Wright and Heikki Nyman, eds. Trans. C. Grant Luckhardt and Maximilian A. E. Aue. Oxford: Blackwell, 1980.

—. *Zettel.* G. E. M. Anscombe and G. H. von Wright, eds. Trans. Anscombe. Oxford: Blackwell, 1967.