The Social Re-Construction of Agency

Katarzyna Paprzycka

Abstract. One of the deep roots of opposition to social constructionism is the belief that the very idea of a *social* construction of physical concepts is highly suspect. In this paper, I want to call attention to the fact that such "constructions" can occur in the opposite direction as well. According to responsibilism, attributions of actions are to be understood in terms of ascriptions of responsibility. Responsibilists thus take the notion of action to be a social concept. I point out how, from the responsibilist point of view, the concept of action is misconstrued as mental by the predominant intentionalist approach in philosophy of action.

One of the deep roots of opposition to social constructionism is the belief that the very idea of a *social* construction of physical concepts is highly suspect. In this paper, I want to call attention to the fact that such "constructions" can occur in the other direction, as it were, as well. There are concepts that really are social and yet we have a tendency to construe them as mental or psychological concepts.[1] In this connection, I want to call attention to a responsibilist theory of agency. According to responsibilism (of which H.L.A. Hart's ascriptivism is an example), actions do not exist as events in the world. Rather, we should understand the attributions of actions in terms of our holding one another responsible for certain events in the world. One of the advantages of responsibilism is that it can capture rather easily a wide class of forms of agency such as omissions (including unintentional omissions), spontaneous, arational, habitual, automatic actions, as well as slips and mistakes. By contrast, the intentionalist approach to agency, according to which an action is paradigmatically a bodily movement caused in the right way by (or otherwise appropriately related, e.g. by means of teleological relations, to) an appropriate rationalizing mental state (belief-desire, intention, intention-in-action, volition), has to either discount such performances as nonagentive or else has to stretch mental concepts so as to encompass the performances. It is this concept stretching that underlies the "mental construction"

---

[1] In calling attention to such mental constructions of what is at roots a social phenomenon, I am not siding with social constructionism. In fact, the debate between social constructionism and realism frequently is really a debate between conceptualism (or antirealism) and realism (see also Hacking 2000), where the modifier 'social' does not play a great role.

in question. From the responsibilist point of view, intentionalism has a tendency to (mis)construe as primarily mental concepts that are, at roots, social.

In §1, I distinguish various senses of the notion of action and try to make clear the concept I am after. In §2 I sketch the distinction between the two types of theories of action and show them to have a common root in Aristotle's remarks on what is voluntary. In §3 I sketch Hart's version of responsibilism and briefly answer some immediate objections (§4). In §5 I sketch one way of understanding the notion of control that I take to underlie responsibilism. In §6 I point out some advantages of responsibilist theories and show how, from the point of view of responsibilism, intentionalism undertakes a "mental construction" of what are ultimately social phenomena.

1. Action as a Unit of Conduct

Perhaps the most fundamental difficulty in analyzing the concept of action is the fact that it plays a significant role in a number of disciplines as diverse as physics, biology, psychology and sociology. As a result, the concept has coalesced a great variety of intuitions. It is thus important to at least try to distinguish some ways of understanding the term 'action'.

(i) There is a concept of *inanimate* action. When a billiard ball thrusts into another billiard ball it acts on the other. To its action, by Newton's third law, there corresponds an appropriate reaction of the other ball. The application of teleological concepts to inanimate action appears to be only derivative. For example, we can speak of the purpose or function of a piece of a thermostat but its purposefulness is derived from its being designed by someone.

(ii) We can speak of the actions of various parts of animal bodies. This is the first stage at which non-derivative teleological concepts find application. The liver's excreting bile, the heart's pumping are examples of what one might call *purposeful* movements or actions.

(iii) The third level is that of *purposive movement* or action. The subjects of our attributions of purposive movements are no longer parts of bodies but rather agentive systems. The movements a spider produces in spinning a web constitute purposive movements. In this sense also, a drug addict's compulsively taking a shot is purposive (Frankfurt 1988, pp. 76-77). Arguably, sleep-walking, some actions performed under hypnosis, as well as the little movements one performs to alleviate muscle pain in one's sleep are purposive. So are feeding the cat, conversing, looking out of the window, walking through a forest.

(iv) The latter but not the former examples belong to a more restrictive category of *intentional movements*. A movement is intentional just in case there is some description under which it is intentional. The category of intentional movements is an extensional category — it

2

picks out a class of events. As such, it is a very different concept from the concept of intentional action, which is an intensional concept (Anscombe 1957; Davidson 1971). Both intentional and unintentional actions, as they are usually understood, are intentional movements in this sense.[2]

It is not uncontroversial to sharply distinguish the category of intentional movements from the category of purposive movements. One might treat the distinction to be one of degree rather than principle. Yet many of the examples relevant here are at least very different from the examples of purposive movements. So when one goes to a rally, one performs an action of a different sort than if one went there in one's sleep.

(v) The fifth sense of 'action' derives from the idea of an agent's overall conduct. Someone's conduct includes her intentional and unintentional doings but also intentional and unintentional not-doings (omissions). When we inquire after a person's conduct during a rally, say, we will be interested in the things the person said and did as well as the things that he omitted to say or do.

The concept of action as part of an agent's conduct has not been at the forefront of philosophers' concern with agency.[3] Most of the debate has centered around the concept of action in the sense of purposive and/or intentional movement. This is, among other things, because intelligence and reason are most clearly manifested in our acting intentionally. But the philosophical focus on "intelligent agency" should not lead one to think that there is nothing interesting about action but for its rational significance. In fact, there are psychological categories that pertain to our conduct rather than merely to our intentional behavior. The most important among them is the concept of character. Character comprises not only our agentive voice — active intentional rational excursions into the world — but also our idleness, passivity, thoughtlessness, carelessness, forgetfulness — our agentive silence, as it were. The responsibilist is best viewed as trying to capture the fifth sense of the concept of action.

2. Two Types of Action Theories

Two general types of theories of action could be seen to have a source in Aristotle's remarks on the voluntary. Aristotle says:

---

[2] This usage of the term 'intentional movement' is not widespread. The term is introduced by Frankfurt 1988. In view of the dispute between minimalism (e.g. Davidson 1971) and moderationism (e.g. Thomson 1977) in the ontology of action, the claim would have to readjusted. Moderationists would insist that intentional movements are parts (possibly proper parts) of complexes with which intentional (unintentional) actions are identical.

[3] The most obvious exception is H.L.A. Hart (1951) who frequently speaks of the "philosophy of conduct," intending to cover both actions and omissions (including unintentional ones) by the term.

> What comes about by force or because of ignorance seems to be involuntary. What is forced has an external origin, the sort of origin in which the agent or victim contributes nothing — if, e.g. a wind or human beings who control him were to carry him off.[4]

These remarks are suggestive of a certain picture of what it means for a performance to be a mere happening rather than an action:

(N$_e$) The agent's φing was a mere happening (nonaction) iff "external forces" caused him to φ.

and a corresponding picture of what it means for a performance to be an action:

(A$_i$) The agent's φing was an action iff "internal forces" caused him to φ.

(N$_e$) can be thought to capture the central thought of responsibilism while (A$_i$) captures the central thought of intentionalism. The ideas of "internal forces" and "external forces" are treated as a stand-in for a more detailed account.[5]

Despite appearances, however, (N$_e$) and (A$_i$) are not so easily reconcilable with one another. In particular, it is worth stressing that (A$_i$) does *not* follow from (N$_e$), nor (N$_e$) from (A$_i$). In order to establish logical relations between them, one would have to relate the concepts of mere happening and action, on the one hand, and of "external" and "internal forces," on the other.

The former move is the least problematic of the two: we can view the concepts of action and nonaction as complementary (relative to a class of performances[6]):

(A-N) A performance is an action just in case it is not a mere happening (nonaction).

Given (A-N), we can establish that what follows from (N$_e$) is:

(A$_e$) The agent's φing was an action iff it was not caused by "external forces."

Analogically, what follows from (A$_i$) is:

(N$_i$) The agent's φing was a mere happening iff it was not caused by "internal forces."

The question is whether (A$_i$) and (N$_i$) (as well as (N$_e$) and (A$_e$)) can be reconciled. They can provided that the following claim is true:

---

[4] Aristotle, *Nicomachean Ethics*, trans. Terence Irwin (Indianapolis: Hackett, 1985), 1110a1-4.

[5] One must remember to avoid simple-minded interpretations here. The distinction is not (as suggested by the form of words Aristotle sometimes uses) between forces outside and inside the agent, for there can be the wrong kind of forces inside the agent (spasms, e.g.), and there may be the right kind of "external" forces (e.g. when someone helps an old person through the street). See also (Frankfurt 1988).

[6] The use of the term 'performance' is theoretical. It is constructed in such a way as to encompass both actions and nonactions – the winkings and the blinkings, the arm raisings and the arm risings, falling off the stairs and running down the stairs etc.

(i-e) "internal forces" caused α to φ iff "external forces" did not cause α to φ.

The status of claim (i-e) has not been given enough attention. What is to preclude the possibility that both types of causes occur at the same time (see Paprzycka forthcoming)? Traditionally, philosophers of action have adopted two strategies. One might begin with elucidating the idea of what it means for the "internal forces" to cause an agent's performance (A_i) and then explain what it means for the agent's performance to be a mere happening in terms of (N_i). This strategy is typical of intentionalism. What this special origin is will depend on the theory in question. The causal theory of action could be seen as paradigmatic for this approach, and one which has reigned over philosophers' intuitions about action for a very long time. On that view, an action is an event that is caused in the right way by appropriate mental states (e.g. Brand 1984; Davidson 1963; 1973; Mele 1992; 2003; Searle 1983). But even many of the challengers to the causal theory share the basic intuition. On the ever more popular agent-causal views (e.g. Chisholm 1976; Lowe 2008; O'Connor 2000), an action is an event that also has a special origin – it must be caused by an agent. Volitionalist approaches (e.g. Ginet 1990) likewise follow this general line of thinking.

Alternatively, one might begin with the idea of what it means for "external forces" to cause an agent's performance (N_e) and then explain what it means for the agent to act by appealing to the absence of such forces (A_e). This is the strategy pursued by responsibilism. From its point of view, the idea of "internal forces" causing the performances is a hypostatization of the absence of such causation by "external forces." In other words, responsibilists take (i-e) to be providing a reductive definition of what it means for the "internal forces" to be in operation. The proper theoretical work is done by the notion of the absence of "external forces." This kind of approach to action has been proposed by H.L.A. Hart (1951), and unfortunately much forgotten since (though see Paprzycka 1997; 2008; Sneddon 2006; Stoecker 2001; 2007). On Hart's view, the notion of action is best understood as a complement to the notion of nonaction (of a mere happening). Nonactions in turn are understood in terms of the presence of defeating conditions (e.g. spasms, ticks, etc.).

What is characteristic of responsibilism is that we are entitled to attribute actions to one another by default, as it were, but we are committed to withdrawing the attribution if certain defeating conditions are present. On the intentionalist approach, by contrast, we are entitled to attribute actions to one another only if we have reasons to believe that the event has had the appropriate origin.

3. H.L.A. Hart's Theory of Action

Hart argued that the concept of action, like the concept of property, is essentially normative and social in that it presupposes accepted rules of conduct. Just as it does not make sense to think that a statement like "Smith owns this piece of land" is a sentence that is "concerned wholly with an individual" (Hart 1951, p. 161), so statements like "Smith did it" likewise should not be understood as describing one individual only. The main purpose of action statements is to attribute ("ascribe" in Hart's language) responsibility for certain events[7] to individuals on the basis of generally accepted rules of conduct.

Hart proposes that claims like "John broke the glass" not be interpreted as describing an action but rather as ascribing responsibility to the agent (here: for the glass breaking). Action claims are ascriptive rather than descriptive. They are never true or false; they may only be appropriate or inappropriate in view of relevant conditions. Transposed from the formal into the material mode, there are no actions among the ontological furniture of the world.

What distinguishes actions from mere happenings, on Hart's view, is not any ontological fact but rather the appropriateness of ascribing responsibility for events in certain conditions (when we intuitively think of them as actions) and the inappropriateness of ascribing responsibility for events in other conditions (when we intuitively think of them as mere happenings). This is what it means to say that the distinction between actions and mere happenings is normative in nature. But this is not yet to give an account of the distinction. In fact, Hart never provides such an account but rather notes that there are conditions that determine whether it is appropriate or inappropriate to ascribe responsibility to the agent.

The structure of action attribution is characteristically defeasible. First, there are, in Hart's terminology, positive conditions that establish the prima facie applicability of the responsibility attribution. In our example, such conditions include John's arm moving in such a way as to break the glass. Second, there are negative (defeating) conditions that defeat the prima facie appropriateness of ascribing responsibility to the agent. Such conditions include John's arm moving because of a spasm, say, or John's arm moving as a result of someone, Mary say, taking it and guiding through the motion. When we learn that it was Mary who took John's arm and

---

[7] It is not exactly clear on Hart's account what ontological category the variable x ranges over in the expression 'α is responsible for x'. Some critics (e.g. Pitcher 1960) have charged Hart with the view that the variable ranges over actions, thus rendering Hart's account circular. But, most of the time (except for a noncommittal statement on the first page of his paper), Hart is quite careful not to talk this way. The variable could be interpreted as ranging over events or events under a description. (For details, see Paprzycka 1997, ch. 2. See also Sneddon 2006.)

smashed the glass with it, we would no longer attribute responsibility to John for breaking the glass. Mary's taking John's arm counts as a defeating condition.

It should be pointed out, however, that defeating conditions can themselves be defeated. Suppose that John held Mary at a gunpoint and ordered her to take his arm and break the glass with his arm. Such a condition defeats the original defeating condition. As a result, it is reasonable to hold John to be responsible for the glass breaking after all. He did it in an unusual way but he did break the glass.

This structure allows us to understand the difference between actions and mere happenings or between it being appropriate and it being inappropriate to ascribe responsibility to an agent. It will be inappropriate to ascribe responsibility to the agent if either no positive conditions are present or while the positive conditions are present some (undefeated) defeating condition occurs. It will be appropriate to ascribe responsibility to the agent if the positive conditions occur and no (undefeated) defeating conditions are present.

4. Some Objections

Many objections could be raised to such an approach. Let me consider three.

First, what is one to do with *mundane actions*. Even if John may be responsible for breaking a vase by raising his arm, what if nothing of any (moral or legal) significance happened as a result of John's movements. He just raised his arm. Likewise, he just sang in the shower. Surely, we want to say that these are John's actions but it seems a stretch to think that he is responsible for them.

A responsibilist can think that John is neither legally nor morally responsible for singing in the shower. There are a couple of options open here. One would be to say that the concept of action is built over the concept of responsibility but applies more widely than does the former. We speak of actions not just when we hold agents responsible but when we could hold them responsible, i.e. when it would be appropriate to hold them responsible. It does not matter that nothing of substance hangs on John's singing in the shower as long as *if it did*, it would be appropriate to hold John responsible. Another sort of response would be to propose a concept of the right sort of responsibility, which identifies those elements in the concept of responsibility that are relevant to the attribution of action. This is the path I have followed in (1997) by proposing the concept of practical responsibility (briefly sketched in §5).

Second, actions are usually understood as performances that are intentional under some description (Anscombe-Davidson thesis), yet the idea of *intentional action* seems completely lost in responsibilism. It is true that the idea of intentional action loses its center-stage role. There is

thus no conceptual pressure for the responsibilist to uphold Anscombe-Davidson's claim. However, this is actually an advantage of the stance rather than its disadvantage.[8] One of the notorious features of intentionalist views is that because they are under pressure to uphold Anscombe-Davidson's thesis, they are also under pressure to stretch the concept of intentional action in ways that we might otherwise resist.

On the other hand, although the concept of intentional action does not ground the concept of action on responsibilist views, there remains a need to understand what intentional actions are. Conceivably, a responsibilist could even adopt a causalist account of intentional action. The main thrust of responsibilism is to deny the claim that the notion of intentional action is conceptually prior to the notion of action. The responsibilist claims that the reverse is true.

Third, there seems to be *no role for beliefs, desires* and other mental states. Again, the responsibilist will say that the expectation that mental states play any role in an account of action is to be rejected with intentionalism. This is not to say that she might not return to them in an account of action explanation. Arguably, however, she is not forced to limit herself to the agent's mental states in an account of action explanation. In fact, the responsibilist may adopt a non-individualist account (see e.g. Baier 1985; Collins 1987; Nowak 1987; 1991; von Wright 1983; Paprzycka 1998; 2002; Wilson 1989), according to which one may very well appeal, for example, to others' mental states in explaining the agent's action (e.g. Susie went to the store because her mother wanted some carrots, Joe shocked the victim because he was told to do so by the experimenter, etc.).

Last but not least, there is what deserves to be called the fundamental problem. Usually we think that we are responsible for what we do. Usually we take this to mean that the notion of action is conceptually prior to the notion of responsibility. We need to settle what was done, first, in order to determine whether we are responsible for it. Yet the responsibilist claims that this logical order should be reversed and proposes to build a notion of action on the basis of the concept of responsibility. It looks like the enterprise is bound to be circular (see e.g. Pitcher 1960). Contrary to appearances this is not a devastating objection. The concept of responsibility is not homogeneous. Particularly helpful in this regard is K. Baier's (1980; 1987) distinction between backward-looking concepts of responsibility (answerability, culpability, liability), which presuppose that the agent has something (possibly an action) to answer for, and forward-looking concepts of responsibility (task-responsibility), which do not presuppose that the agent has done something. In fact, the account of practical responsibility sketched below involves the latter

---

[8] It should be noted that there are ways of developing responsibilism where the notion of intentional action does play quite a significant role (see e.g. R. Stoecker 2001; 2007).

notion. It is beyond doubt, however, that the objection has to inform the responsibilist theorizing and keep him on his toes.

5. Practical Responsibility

The general structure of responsibilist accounts can be perhaps cast in the following way. Our attributions of actions of φing to agents presuppose that we take agents to be practically responsible for φing. Agents are, in turn, practically responsible for φing when they are in control of φing or when φing is within their power. When α is practically responsible for φing and φs, α's φing is an action.

Normally we take ourselves to be – and indeed are – in control of various ordinary activities (raising arms, moving feet, reaching for glasses, pouring milk, running, writing, etc.). Arguably, we differ in the repertoires of activities we are in control of – some of us can juggle, dance or raise arms in graceful ways, others cannot sing, cannot write or may be incapable of walking. However, the very application of the concept of action in our social lives presupposes that the range of activities that we are in control of overlaps quite a bit. The understanding of the idea of what it is to be in control of φing includes at the very least that the agent is capable of reliably fulfilling the task to φ (and the task not to φ).[9]

Defeating conditions – or Hart's negative conditions – are conditions that affect the agent's control of an activity. When someone breaks a leg, she is no longer in control of running or even walking briskly. When someone acquires a tick, he cannot be relied upon to wink as a sign to start a revolt.

The heart of the responsibilist account involves understanding these key concepts, what it is for the agent to be in control of φing and how various conditions can affect this state. Once the responsibilist demonstrates that most agents indeed are in control of φing, where 'φ' ranges over "common" action types, he entitles himself to the characteristic feature of responsibilist theories, viz. that such action types lie within the agentive purview by default. With respect to them, we are *guilty of being agents until pro*

ven innocent, so to speak.[10] According to the responsibilist, one does not need to demonstrate any special origin for a performance to count as an action as long as it is a

---

[9] In a remarkable paper, A. Baier (1971) has sketched an approach to action where the idea of a task plays a central role. The account sketched owes a lot to hers.

[10] The thought that some performances acquire the status of actions (indeed intentional actions) by default is present in Brandom's account (see esp. Brandom 1994, pp. 257ff). While Brandom officially adheres to a causal account of action, this thought makes him a potential responsibilist.

performance of a type that falls within the range of those that are within the agent's power or control. Of course, such action attributions are defeasible – they are sensitive to the occurrence of any circumstances that affect the agent's control of those types of actions.

While it is beyond the scope of this paper to go into the details of responsibilist positions, I want to briefly provide a proposal for an elucidation of the relevant idea of "what it is to be within the agent's power or control" in terms of the idea of reasonable normative expectations. I want to stress, however, that this proposal is not inherently tied to responsibilism. It is one way of developing the position but probably not the only way of doing so.

Two intuitions seem to be central to our idea of what it is to be within the agent's power to do something. First (the success intuition), the agent must be able to succeed in φing. It is not within the power of a one-year old to write a novel, or of someone who broke a leg to run a marathon. Second (the difference intuition), the agent must be able to make a difference – it is not within our power to make the sun shine, or to hold breath for an hour.

One way to capture these intuitions is to envisage an ideal test to which an agent could be subjected. We give the agent a series of tasks, to which he responds in the best possible way; we are assuming, in other words, that he is cooperative, that there are no other designs, intentions, expectations in play, the agent is at ease, under no pressure, etc.[11] The tasks are of two kinds, to φ and not to φ, and they are interspersed randomly in a series.

	Task: φ	Task: not-φ
(i)	pf-fulfilled (φ)	pf-fulfilled (not-φ)
(ii)	pf-fulfilled (φ)	pf-frustrated (φ)
(iii)	pf-frustrated (not-φ)	pf-fulfilled (not-φ)
(iv)	Other	

Table 1. Possible result patterns of a simplified test sequence

Three situations are of special interest. Suppose that an agent systematically pf-frustrates[12] the expectation to φ (situation (iii) in Table 1). When he is expected to φ, he does not.

---

[11] This is an idealizing assumption. I am making it in order to sharpen the intuitions at stake.

[12] The notions of prima-facie (pf-) fulfillment and pf-frustration are introduced in part to prevent the circularity problem that affects responsibilist accounts. One might, for example, argue that an expectation to raise an arm is not fulfilled when the arm rises of its own accord, by accident, in a spasm etc. An account of action that used such a concept of expectation fulfillment would be rightly charged with circularity. That is why, at this stage, I appeal to a very liberal notion of pf-fulfillment that includes not only actions but also nonactions as pf-fulfilling an expectation.

In such a case, it would be unreasonable*[13] to expect of the agent that he φ. The agent cannot succeed in meeting the task. Suppose that the agent regularly pf-fulfills the expectation to φ but also regularly pf-frustrates the expectation not to φ (ii). What this will mean is that the agent φs indiscriminately. In such a case, we would tend to think that the agent's φing is not up to him, that the agent cannot make a difference, and hence that it would be unreasonable* to expect of him that he φ. This configuration would obtain if we expected the agent to breathe, for example. Finally (i), when the agent pf-fulfills all the expectations (when expected to φ, the agent responds by φing, when expected not to φ, the agent responds by not φing), we would tend to think that φing and not φing are "within the agent's power," that it is reasonable* to expect of the agent that he φ.

We can understand the success and difference intuitions accordingly:

Success condition: It is prima facie unreasonable* to hold α to a (normative) expectation to φ if were α subjected to the test (with all of its conditions satisfied[14]), the expectation to φ would be pf-frustrated.

Difference condition: It is prima facie unreasonable to hold α to an expectation to φ if were α subjected to the test, the expectation to φ would be pf-fulfilled but the expectation not to φ would be pf-frustrated.*

It is possible for a normative expectation to be prima facie reasonable* (e.g. to run a race) but for certain conditions to occur (e.g. a broken leg) that would defeat its reasonableness*. A person who was subjected to the test in the presence of such conditions would respond quite differently. We can distinguish hindering and compelling conditions accordingly:

Condition C is a **hindering condition** with respect to an expectation to φ iff were α subjected to the test in condition C, the expectation to φ would be pf-frustrated.

Condition C is a **compelling (or forcing) condition** with respect to an expectation to φ iff were α subjected to the test in condition C, the expectation to φ would be pf-fulfilled but the expectation not to φ would be pf-frustrated.

---

[13] I mark the notion of reasonableness with an asterisk to note that it is a theoretical concept, which captures but one dimension of our ordinary concept. In particular, our concept of what it would be reasonable to expect of another person also includes a normative component (it would be unreasonable of me in this sense to expect of my neighbor to mow my lawn even though it would be within his power to do so), which I am ignoring here.

[14] I take the condition 'α is subjected to the test' to be synonymous with the expression 'α is subjected to the test and α fulfills all the conditions of the test, i.e. is cooperative, at ease, under no pressure, with no other intentions or expectations in play'.

An example of a hindering condition with respect to an expectation to run a race would be breaking a leg. An example of a compelling condition with respect to an expectation to walk would be being forced to do so by another person. We may perhaps distinguish yet a third type of condition, whose occurrence would lead a person to lose control in yet another sense. The person who normally responds reliably to expectations to touch the table and not to touch the table might simply become erratic. Arguably a person with some neural condition or who has taken drugs might fit this kind of pattern. A person in this condition subjected to the test would not exhibit a pattern of responses falling neatly under rows (i)-(iii) but rather fall in row (iv). Such a condition might be called a disabling condition.

An agent α is practically responsible for φing at time t if: (a) the expectation to φ is prima facie reasonable*, (b) at time t, no (undefeated[15]) defeating condition (with respect to the expectation that α φ) is present.

We can use the notion of practical responsibility in the responsibilist account of action. When John's arm rises it will count as his action as long as he is practically responsible for raising his arm at this moment, or as long as it would have been reasonable* to expect of him that he raise his arm at this moment. If John is like most of us then if we were to subject him to the test (as long as its conditions are satisfied, i.e. John is cooperative etc.), his performances would fit the pattern (i). In other words, the expectation of John to raise his arm is prima facie reasonable. At the same time, as long as no undefeated defeating occurs, the expectation to raise an arm continues to be reasonable*. So as long as John is practically responsible for raising his arm (at t) and his arm rises (at t), we can attribute to John the action of raising his arm. The attribution does not depend on our knowledge of what mental state John was in at all.

6. The Social Reconstruction of Agency: The Case of Intention-in-Action

Central to the intentionalist view of agency is the thought that the agents' bodily movements be appropriately related to the agents' reasons (often conceived of as mental states that rationalize the actions). Central to the responsibilist view is the thought that the agent have an appropriate sort of control, the lack of which would exclude responsibility. Responsibilists take competent agents to perform actions by default. No special account needs to be given to count an agent's walking movements, for example, as an action. The agent need not have a reason to walk (though she might), a desire to walk (though she might), an intention to walk (though she might) or

---

[15] Defeating conditions can be defeated by other conditions and those conditions can be defeated by still other conditions.

intention-in-action or volition or … The walking movements of a competent and skilled walker will count as an action of walking whatever its causes and whatever the mental attitude of the walker, as long as the walker is in control, i.e. as long as nothing defeats reasonableness of the expectation of the agent that she walk.[16]

In giving this sort of account of agency, we can give a more accurate account of some types of agency, without having to venture to search for intentional states, the postulation of which often seems just arbitrary. Consider the introduction of the notion of intention-in-action by Searle. He first expresses the belief that there are no actions without intentions (Searle 1983, p. 82). Then he argues that because there are cases of actions which are not preceded by a prior intention, so in such cases the intention must be *in* the action.

In other words, one starts with the theoretical claim that for a performance to be an action, it *must* be suitably related to an intention. But there are actions which are not done on (prior) intentions. In fact, there are spontaneous actions, actions done for no reason (like one's humming a tune under shower). Rather than taking them as a problem for the theory, one saves the theory by postulating that in case of those performances there are intentions after all, they are just concurrent with actions — intentions-in-action. Thus introduced the concept of intention-in-action is an *ad hoc* device introduced to save the theory that actions must be related to intentions. Moreover, we may suspect that the concept of intention-in-action is really parasitic on the concept of action. We have a much firmer grip on the concept of action than we do on the concept of intention-in-action. To the extent that the notion of intention-in-action serves as an *explanation* of what an action is, it is based on theoretical trickery. We have first coined the notion of intention-in-action on the basic of our intuitions about the concept of action. Then we use the conjured concept to explain our concept of action.

Searle uses the following example to make his suggestion intuitive:

> Suppose you ask me "When you suddenly hit that man, did you first form the intention to hit him?" My answer might be, "No, I just hit him." But even in such a case I hit him intentionally and my action was done with the intention of hitting him. I want to say about such a case that the intention was in the action but that there was no prior intention. (p. 84).

It would indeed be unreasonable to be suspicious of the possibility that an intention may arise on the spur of the moment. But to allow such a possibility is not yet to buy into the claim that an

---

[16] For similar reasons, the responsibilist is open to the idea of letting in as actions cases that are notoriously problematic for the intentionalist because the requisite internal make up seems lacking: arational actions (Hursthouse 1991), habitual actions (Pollard 2003; 2006), nonintentional actions (Chan 1995), unintentional omissions (Smith 1984; 1990), mistakes and slips (Peabody 2005). At the same time, cases of antecedential wayward causal chains do not present any special problems (see also Paprzycka 1997; forthcoming).

intention in action, understood as a mental state, is present in all cases of action. Consider the following case.

Suppose that I walk down the street, engrossed in thoughts, picking leaves from nearby bushes as I walk by them. You catch me doing this. I may have not even realized that I was doing so. You ask "When you picked those leaves, did you form an intention to do so?". I am very likely to answer just like Searle did – "No. I was just picking them, I suppose." Did I do so *with the intention* of picking the leaves? This is a very strange question to ask in this situation. If what you mean by that is "Was I picking the leaves?" then my answer will be positive. (It is clear, however, that you are not asking this question since you knew I was picking the leaves before I did.) If you are asking about my intention then I would be inclined to answer in the negative. I certainly had *no reason* at all to pick them. I would also oppose the view that I wanted to pick them (Paprzycka 1998; 2002).

Perhaps the intention is just the goal toward which my movements were directed. However, consider a case of an action slip reported by James:

> Very absent-minded persons in going to their bedroom to dress for dinner have been known to take off one garment after another and finally to get into bed, merely because that was the habitual issue of the first few movements when performed at a later hour. (James 1890/1983, 119)

What are we to say about such a case? The person intends to dress for dinner and thus *intentionally* prepares to go to bed? This is not a case of a change of mind: the person still intends to prepare for dinner and goes on to do so after realizing the mistake. This is not a case of a mistaken belief: the person does not have the belief that going to bed is a good way of dressing for dinner (see also Peabody 2005). Does the person have an intention in action to go to bed? This is what the person is doing. He has no reason to do it. In fact he has reasons not to do it. His movements are directed at going to bed as they are when he goes to bed with the intention of doing so. The problem is that unlike in the usual case, in this case the agent disowns that goal. If he attends to the situation, he will vehemently deny that this was his intention.

Note too that this is to allow that when one looks at the slip from a psychological or neurocomputational point of view one will be able to find the activation of states that are responsible for certain motor routines (sometimes also referred to as "motor intentions"). But there is a gulf between such states and philosophers' intentions, which rationalize the action and which are responsible for the action's intentional character. The neurocomputational difference between what happens in the case when someone intentionally dresses for dinner and when someone intends to dress for dinner and slips into preparing for bed may be a matter of degree (perhaps the higher activation of one or two units made the difference). From a philosophical

point of view the difference is qualitative – in one case the person acts rationally and intentionally, in the other – the performance is irrational in these circumstances and not intentional.

From a responsibilist point of view, this tension between the notion of intention, which is tied with the rationalization of an action, and the notion of intention, which is responsible for a performance being an action, can be removed. The notion of intention can be reserved for giving an account of the rationalization and intentionality of action. There are good reasons to suppose, however, that not all of our actions are performed for a reason. There are good reasons to suppose that not all of our actions are intentional under some description. All of our actions presuppose, however, that we exhibit the right sort of control, which is prerequisite of our being responsible.

If this account is correct then the intentionalist presents us with a "mental construction" (here: the postulation of intention-in-action) of what is at roots a social phenomenon (here: the absence of defeating conditions). This is indeed one of the complaints of H.L.A. Hart who warned against this kind of construction:

> These positive-looking words 'intention', etc., if put forward as necessary conditions of all action only succeed in posing as this if in fact they are a comprehensive and misleadingly positive-sounding reference to the absence of one or more of the defences, and are thus only understandable when interpreted in the light of the defences, and not vice versa. Again, when we are ascribing an action to a person, the question whether a psychological 'event' occurred does not come up in this suggested positive form at all, but in the form of an inquiry as to whether any of these extenuating defences cover the case. (Hart 1951, p. 163).

The responsibilist advocates a view of agency, which is fundamentally social. It is only available within quite complex practices of holding one another responsible. In resisting the idea that agency is a social concept, we have become all too accustomed to thinking that it is a mental concept. As a result, we have found it much easier to postulate all kinds of mental states or processes than trying to find resources already at our disposal when we view it as a social concept. Perhaps it is time to pause and reflect lest we should witness the debate between realism and physical constructionism in the future.[17]

Bibliography

Baier, A.C. (1970) "The Search for Basic Actions", *American Philosophical Quarterly* 8: 161-170.

Baier, A.C. (1985) *Postures of the Mind: Essays on Mind and Morals.* Minneapolis: University of Minnesota Press.

Baier, K. (1980) "Responsibility and Action" in: M. Bradie and M. Brand, eds., *Action and Responsibility.* Bowling Green, OH: Bowling Green University Press, pp. 100-116.

Baier, K. (1987) "Moral and Legal Responsibility", in: M. Siegler, S. Toulmin, F. E. Zimring and K.F. Schaffner, eds., *Medical Innovation and Bad Outcomes*. Ann Arbor, MI: Health Administration Press, pp. 101-129.

Brand, M. (1984) *Intending and Action. Toward a Naturalized Action Theory*. Cambridge, MA: The MIT Press.

Brandom, R. (1994) *Making It Explicit*. Cambridge, MA: Harvard University Press.

Chan, D.K. (1995) "Nonintentional Actions", *American Philosophical Quarterly* 32: 139-152.

Chisholm, R.M. (1976) *Person and Object*. La Salle, IL: Open Court.

Collins, A.W. (1987) *The Nature of Mental Things*. Notre Dame, IN: University of Notre Dame Press.

Davidson, D. (1963) "Actions, Reasons, and Causes", reprinted in: *Essays on Actions and Events*. Oxford: Clarendon Press, 1980, pp. 3-19.

Davidson, D. (1971) "Agency", reprinted in: *Essays on Actions and Events*. Oxford: Clarendon Press, 1980, pp. 43-61.

Davidson, D. (1973) "Freedom to Act", reprinted in: *Essays on Actions and Events*. Oxford: Clarendon Press, 1980, 63-81.

Frankfurt, H. (1988) "The Problem of Action", in: *The Importance of What We Care About*. Cambridge: Cambridge University Press, pp. 69-79.

Geach, P.T. (1972) "Ascriptivism", in *Logic Matters*. Berkeley: University of California Press, pp. 250-254.

Ginet, C. (1990) *On Action*. Cambridge: Cambridge University Press.

Hacking, I. (2000) *The Social Construction of What*. Cambridge, MA: Harvard University Press.

Hart, H.L.A. (1951) "The Ascription of Responsibility and Rights", in: A. Flew (ed.), *Essays on Logic and Language*. Oxford: Blackwell, pp. 145-166.

Hursthouse, R. (1991) "Arational Actions", *Journal of Philosophy* 88: 57-68.

James, W. (1890/1983) *The Principles of Psychology*. Cambridge, MA: Harvard University Press.

Lowe, E.J. (2008) *Personal Agency. The Metaphysics of Mind and Action*. Oxford: Oxford University Press.

Mele, A.R. (1992) *Springs of Action*. Oxford: Oxford University Press.

Mele, A.R. (2003) *Motivation and Agency*. Oxford: Oxford University Press.

Mele, A.R. and Moser, P.K. (1994) "Intentional Action", *Nous* 28: 39-68.

Nowak, L. (1987) "Man and People", *Social Theory and Practice* 14: 1-17.

Nowak, L. (1991) *Power and Civil Society: Toward a Dynamic Theory of Real Socialism*. New York: Greenwood Press.

O'Connor, T. (2000) *Persons and Causes. The Metaphysics of Free Will*. Oxford: Oxford University Press.

Paprzycka, K. (1997) *The Social Anatomy of Action: Toward a Responsibility-Based Account of Agency*. Ph.D. Dissertation: University of Pittsburgh.

Paprzycka, K. (1998) "Collectivism on the Horizon: A Challenge to Pettit's Critique of Collectivism", *Australasian Journal of Philosophy* 76: 165-181.

Paprzycka, K. (1999) "Normative Expectations, Intentions and Beliefs", *The Southern Journal of Philosophy* 37-4: 629-652.

Paprzycka, K. (2002) "False Consciousness of Intentional Psychology", *Philosophical Psychology* 15: 271-295.

Paprzycka, K. (2008) "Sneddon on Action and Responsibility", *Polish Journal of Philosophy* 2: 69-88.

Paprzycka, K. (forthcoming) "Can a Spasm Cause an Action?", *Grazer Philosophische Studien*.

Peabody, K. (2005) "Trying Slips: Can Davidson and Hornsby Account for Mistakes and Slips?", *Philosophia* 35: 173-216.

Pitcher, G. (1960) "Hart on Action and Responsibility". *The Philosophical Review* 69: 226-235.

Pollard, B. (2003) "Can Virtuous Actions Be both Habitual and Rational", *Ethical Theory and Moral Practice* 6, 411-425.

Pollard, B. (2006) "Explaining Actions with Habits", *American Philosophical Quarterly* 43: 57-68.

Searle, J.R. (1983) *Intentionality. An Essay in the Philosophy of Mind.* Cambridge: Cambridge University Press.

Smith (Milanich), P.G. (1984) "Allowing, Refraining, and Failing. The Structure of Omissions", *Philosophical Studies* 45: 57-67.

Smith, P.G. (1990) "Contemplating Failure: The Importance of Unconscious Omission", *Philosophical Studies* 59: 159-176.

Sneddon, A. (2006) *Action and Responsibility.* Dordrecht: Springer Academic Press.

Stoecker, R. (2001) "Agents in Action", *Grazer Philosophische Studien* 61: 21-42.

Stoecker, R. (2007) "Action and Responsibility - A Second Look at Ascriptivism", in: C. Lumer and S. Nannini, eds., *Intentionality, Deliberation and Autonomy: The Action-Theoretic Basis of Practical Philosophy*. Aldershot: Ashgate, pp. 35-46.

Thomson, J.J. (1977) *Acts and Other Events.* Ithaca, NY: Cornell University Press.

Wilson, G.M. (1989) *The Intentionality of Human Action.* Stanford: Stanford University Press.