

Collective Responsibilities of Random Collections: Plural Self-Awareness among Strangers

Hans Bernhard Schmid

There are many things people should do—and some of those are things people can only do jointly, together with others. A well-organized society usually has corporate agencies for these tasks. Since crime intervention in larger communities is too big a task for a single sheriff, we have our police departments. To fight fires that nobody can put out by him- or herself there is the fire department. The ambulance is in charge of medical emergencies, for poverty and deprivation there are the social services, for moral catastrophes in remote corners of the world we have foreign aid, and so on. These and similar organizations are the agents to take care of those morally sensitive tasks which no one can perform alone.

We hold these organizations responsible if they fail. In recent philosophical research, it has come to be widely accepted that we are justified in thinking of such organizations as collective agents of their own, and thus as collective bearers of responsibility (cf. List and Pettit 2011). One of the arguments for collective or corporate responsibility is that the responsibility in question often cannot be straightforwardly reduced to the responsibility of the individual members of the organization. Though there are usually individual members to blame, blaming them for failing to do their job, for not performing in their roles, is different from blaming them for the act or omission in question. The chief of the fire department is responsible for failing to re-organize his or her department, and is thus certainly heavily implicated in the blame for the fire department's poor performance. But his or her responsibility is for leading the fire department, not for doing the fire department's job, and similarly for all the other roles in an organization. It thus makes sense to assume that organizations do have responsibilities of their own.

To live in a well-organized society with functioning corporate agents makes it easier for us, as individuals, to live up to our moral responsibilities. If the morally required task at hand is too big for us, individually, all we have to do is basically to dial the right number—call in the police, the ambulance, the firefighters, or whatever other corporate agent is suitable for the job. The moral responsibilities for many of those actions which we can only perform jointly thus comes to be placed on the broad (if somewhat metaphorical) shoulders of our corporate agents.

And yet, even in near ideal circumstances, this is not always enough. Even if the police, the ambulance, and so on are well organized, they cannot be everywhere, or not right away. In such situations, it can be morally required of us, the individuals, to *team up*, join our forces, get our joint effort together and organize ourselves quite spontaneously to do the right thing together, at least until the relevant corporate agent appears on the scene—even if we’re not friends who are already used to deliberate about what to do, how to do it, and to act together. In some situations, there is a moral responsibility to team up with perfect strangers—people who happen to be around simply by chance. This is what has come to be discussed under the title “the responsibility of random collections.”

What kind of responsibility is at stake here? As this seems to be somewhere between the two cases, one way to approach this question is from the distinction between the responsibility of individuals and the responsibility of corporate agents. Thus conceived, the question is this: Is our duty to team up and act jointly with strangers more like an extension of our individual responsibility to dial the right number, or is it more like the responsibility of a make-shift or substitute “citizen’s police’s” responsibility to do their job? The difference is that in the first view, the responsibility involved seems to be of the individual kind, whereas in the second case, it seems to be a collective’s. The strongest intuitive argument against the latter view is that in order to have responsibility, there has to be an agent whose responsibility it is, and that while there certainly are organized group agents, “random collections” are not of that sort. Or, put differently: for somebody or something to be morally responsible, that somebody or something must be, under suitable circumstances, an adequate target of moral blame; yet if people fail to team up to come to your aid, it might seem you would not rationally blame “the bystanders,” collectively, but rather each of the individual bystanders, individually. Random collections, it seems, cannot be the target of moral blame, because they are not agents.

Yet there are problems with the first alternative, too. The strongest intuitive argument is that where strangers fail to team up and do what’s morally required of them to do together, in joint action, it seems what they are responsible for is the *failure to act together*—and this is not exactly the same as the participants’ individual *failure to initiate a team action or to join in*. Individuals can be blamed for not *participating* in a joint action, but the blame for *not acting* is on them, collectively. Thus it seems that there is a way in which such cases involve a responsibility that is the collective’s after all.

The main thrust in the received literature on the responsibility of random collections is to accommodate some of the latter intuition, while remaining firmly committed to the view that whatever responsibility random collections might have cannot be genuinely *collective*, but has to be some *distribution of individual responsibility* (I). In this view, a bunch of strangers simply isn’t the kind of entity that is a suitable target for blame. A random collection of strangers is no collective agent, especially where the random collection fails to team up. Yet I shall argue that the distributive view of the moral responsibility of

random collections makes it hard to give an adequate understanding of how the responsibility in question is not just to *team up*, but rather to *act*, too. I will address an attempt in the recent literature to solve this dilemma by separating collective responsibility from collective agency (II) and develop an alternative view in which there is a condition under which the collective responsibility of random collections is a plural agent's. Where this condition is met, random collection of individuals can be the kind of entity that is a suitable target for blame even where they failed to team up. The condition in question is the capacity for plural self-determination, that is, the right sort of practical knowledge of what the strangers should be doing together (III).

I

Here is a version of the classical example for the moral responsibility of random collections (the original is in Virginia Held's 1970 paper by the title "Can a Random Collection of Individuals Be Morally Responsible?"). The case is a random collection of individuals in a sparsely occupied subway car that witnesses a bully beating up a defenseless person in such a way that it is obvious that the victim's health is in danger. Assume further that—in contrast to Held's example—the bully is obviously much stronger than any one single individual in the random collection of passengers, even if that individual could count on the victim's active support.¹ It is obvious that if any one of the passengers were to intervene alone in this situation, and none of the others were to join in, the helper's health would be in danger, too. However, if at least two of the witnesses would intervene in a minimally organized (coordinated) way, it seems obvious that the abuse could be stopped. Imagine now that no one takes the initiative, so that joint action comes about. What is the responsibility involved in this case? Assuming that not living up to one's moral responsibility is blameworthy, a good point to start answering this question is by asking ourselves: Whom would we blame for what exactly?

Our Held-style example is such that given the risk to any single helper's health, no one can be blamed for not intervening in the victim's behalf alone. It is equally uncontroversial that we would blame each of the passengers for failing to take the initiative, for not trying to gather a group together (e.g., by suggesting a coordinated intervention, etc.). But how can we accommodate the intuition what the fellow passengers failed to do is not only to team up, but *to act*? How much room is there to blame the random collection as such for not intervening, even though it is not an organized group?

Virginia Held argues that "we would hold the random collection morally responsible for its failure to act as a group" (477). This accommodates the intuition that what the random collection failed to do is *to act*, rather than just to team up. However, Held then goes on emphasizing the distributive nature of this responsibility: "if random collection R is morally responsible for the failure to

do A, then every member of R is morally responsible for the failure to do A, although, perhaps, in significantly different proportions” (ibid.).

This view seems problematic (and indeed untenable) for two reasons.

1. The first reason has to do with the difficulty of individuating action A.
 - (a) Is A adequately described as “coming to the victim’s aid”? It seems obvious that this cannot be right. On this reading, saying that the collection is responsible would amount to holding each individual responsible for something he or she is not morally required to do (because of the risk to any single helper’s health), which is against the premise built into the example.
 - (b) Is A perhaps adequately described as “subduing the bully”? This does not seem right either, because on this reading, we would hold each individual responsible for something that individual is not even *able* to do. This raises issues concerning “ought implies can” (which certainly has some intuitive appeal), particularly where *omission* rather than *action* is the target of blame. Where action is concerned, however, parts of the literature on collective responsibility take a different line. As an example from the literature, consider the case of a group of teenagers cooperating in pushing a heavy boulder off a plateau, “so that it rolls down a slope and smashes a car at the bottom” (Zimmerman 1985, 116). Zimmerman argues that in this case, each of the ten participants is “fully responsible” for smashing the car. Taking this to be our case, the view is that if our subway car random collection of individuals were to be blamed for cooperating in beating up the victim rather than for failing to come to the victim’s aid, each member should be held “fully responsible.” Yet even if we accept this harsh line, it seems that it does not translate to responsibility for omission—the “action-omission-effect” that is well-studied in experimental social psychology for the individual case (cf. Bostyn and Roets 2016) thus has interesting ramifications in the case of joint action/omission.
 - (c) If it is not the case that every member of R can be held responsible either for “coming to the victim’s aid” or for “subduing the bully,” a third alternative to individuate action A that comes to mind is that A might simply be something like “acting so as to make it happen that the random collection cooperates in coming to the victim’s aid.” This does not work smoothly with regard to Held’s phrase—the random collection is not responsible for failing to “act so as to make it happen that the random collection cooperates in coming to the victim’s aid,” only the individuals are, and thus this way of individuating A fails to capture the sense in which the responsibility is *the*

collection's. Yet Held seems to give this third reading some support, as she says herself that “it may well follow that in some cases all the individuals in a random collection are responsible for not acting to transform the collection into an organized group, even though none of these individuals is responsible for not taking the action that ought to have been taken by an organized group in these circumstances” (480). Yet “acting to transform the collection into an organized group” seems to be something for which individuals are responsible, and not the collection as such.

It thus seems that none of the three ways to individuate action A is convincing; a) and b) are too strong, c) is too weak. Version (a) blames agents for having failed to do something they are not morally required to do, version (b) blames agents for having failed to do something they are unable to do, version (c) seems reasonable in and by itself, but it does not give us an account of the responsibility we’re trying to explain; it does not amount to “holding the random collection responsible for its failure to act as a group.”

One might think that in spite of all appearances, Held must have a fourth version to individuate “action A” in mind rather than version (c)—a version we have not considered so far. But what could this possibly be? We have considered the most obvious candidates, and found them wanting.

2. The second objection against Held’s distributive reading is that where the random collection of individuals fails to act, but one of its members did his or her best to initiate a joint action, it seems implausible to assume that “each” of the individual members is responsible (cf. Bates 1971). Imagine the following version of the subway scenario from a first-personal perspective. You are one of the passengers. You immediately call the police. But you know that you have to do more. Given the strength of the bully, you’re certainly not morally required to start an intervention by addressing the bully, hoping for another witness to follow suit (this is too dangerous). Rather, what you should do is to turn to your fellow travelers, establish joint readiness and agree on intervening together—perhaps by saying “Let’s stop this, shall we?,” and wait for a nod from your prospective partner—and then do it, together with your partner. Imagine that for some reason, you just can’t find any partner. You look around, but the other people simply avoid making eye contact with you. You call on them, perhaps by saying “we must do something—will anyone help me to stop this, please?,” but to no avail. You do whatever you can to team up with some other witness to stop this, without putting your life in danger. But you fail. You just can’t find any partner for an intervention. At the next subway station, the police come and subdue the bully—but by that time, the victim has suffered further harm that could have been prevented by some decisive team action. Certainly, this is good reason for blame.

The point of this version of the example is that it is still true, in this case, that the random collection failed to act as a group in this scenario. But how exactly can we make sense of the intuition that “we would hold the random collection morally responsible for its failure to act as a group” here? The distributive reading does not seem to work smoothly, given your individual efforts to initiate a rescue mission. It is not the case that the responsibility can be distributed to *all* the individual members: after all, you did whatever you could to transform the collection into an organized group. It is not that you just have a lesser share in the distribution of individual blame. Rather, it seems that you, *personally*, are not to be blamed at all. This seems to leave us with three options. Either the collection is not responsible after all (i), or the collection is responsible, but you’re not part of it (ii), or the collection is responsible, but not in the distributive way sketched by Held (iii).

Let’s start with (ii). Bates (1971, 347) dismisses this move as “arbitrary and ad hoc,” yet in our scenario, it seems natural for you to say that “they,” the random collection consisting of the other passengers (excluding *you*), are morally responsible. Yet what this collection of individuals is responsible for is clearly not “its failure to act as a group.” It’s not the collection of the witnesses *minus you* that failed to act as a group, but rather the collection of the witnesses *including you* that failed to act as a group. The failing collection includes you, too. It is not the case that the collection should have acted jointly *without* you. Rather, the members should have acted jointly *with* you. It is for this reason, and not because the move is simply arbitrary, that holding the collection responsible, but excluding you from the collection, is not plausible.

This leaves us with the option of letting go of the idea of the collection’s responsibility altogether and biting the bullet on the responsibility gap (i), or assuming that the collection’s responsibility is of the collective rather than the distributive sort (iii). Rather than arguing against the former alternative, I’ll focus on the latter option in the following.

II

In the core sense of the word at least, responsibility is for action (including omission), and at least as far as moral theory is concerned, it is more often a matter of blame rather than praise. Where we hold people responsible for events or states of affairs, we do so in virtue of some wrongdoing or wrongful omission. In blame, we are thus addressing agents. Furthermore, we hold them responsible for *their own* actions. Random collections, however, do not seem to qualify as agents. None of the criteria listed in the rich recent literature on group agency seem to apply to random collections. There is no joint commitment of the sort Margaret Gilbert (1989) assumes in her account; there is, it seems, no “rational unification” of a single perspective required by Carol Rovane (2004), and certainly there is no established decision procedure of the sort analyzed by Christian List and Philip Pettit (2011).² Therefore, it seems that random

collections such as our passengers in the subway car simply are not a suitable target of blame, because they are not (group, collective, or plural) agents.

One way to go from here could be simply to detach collective responsibility from group agency. Sara Rachel Chant has made this suggestion in a recent paper (2015). Her fictional case is what she now calls a Hollywood Standoff (formerly known under the politically a bit less correct label “Mexican Stand-off”). This is a situation in which three agents hold each other at gunpoint in such a way that A threatens B, B threatens C, and C threatens A. The first to shoot will surely be shot, and if anyone lowers his or her gun, chaos would ensue, anyone could be shot, and it is likely that somebody will be shot. Chant, in her scenario, has a child dying of thirst within sight. In order to make this a clear case of responsibility for joint action, Chant’s example assumes that the three Hollywood standoffers could save the child only by each one giving the child the last sip of water from his or her flask, so that everyone should contribute. In this situation, Chant argues, we have a pure case of collective responsibility without distributive personal responsibilities—no one in the collection of three is responsible for not lowering his gun and coming to the child’s help (which he or she is not morally required to do because of the risk to his or her life), let alone for the death of the child (which he or she could not prevent with the amount of water he or she has at her own disposal). But as a collection, they are still responsible for the child’s death. Chant argues that this thought experiment shows that there is collective, nondistributive responsibility without collective agency, as the three parties in the standoff are not a collective agent.

Chant’s thought experiment is striking in that it clearly exempts each participant individual from individual responsibility, while supporting the intuition that there is something *they*, together, should do. However, the central question concerning collective responsibility is basically the same in Chant’s and Held’s scenarios, and it does not seem to be addressed in Chant’s short paper: Even if we accept, for the time being, the view that the random collections in Held’s and Chant’s examples are not collective agents, we need to answer the question: *in virtue of what* is the collection’s failure to do the right thing *the collection’s own omission*? The intuitive answer seems obvious: it is because these three people, together, choose to threaten each other rather than to come to the child’s help—they, together, could help, and they, together, know it—so whatever else ensues is *their* omission. Perhaps we can attribute responsibility to entities that are not of the sort analyzed in current philosophical research on group agency; but surely, if we understand responsibility as related to praise and blame, there has to be a way in which the failure in question is still the collection’s *own* doing. If we want to hold on to the intuition that such collections are collectively responsible, we need an account of how the doing in question is the collection’s.

This seem to exert some pressure on seeing the random collections in our examples as *some sort of agents* after all—as entities of whom it makes sense to say that some action or omission is *their own*. In her treatment of the issue,

however, Chant argues that the claim that, under certain circumstances at least, random collections can be agents, “is extremely difficult to sustain”:

To argue that the group is a collective agent, it is necessary to establish that the group has some of the relevant features of agency. By stipulation, the group has none of those relevant qualities. It does not have a decision procedure, nor does it act with any collective purpose or seek any specific collective goal. (Chant 2015, 89)

In a footnote to this passage, Chant refers to the received accounts of group agency quoted above; and yet, that random collections of our kind do not appear to be agents, according to these views, this might just be these views’ mistake. Put briefly and polemically, defining agency by decision procedure or purpose is capturing agency by its (normal) *effects* rather than by its principle. It is simply putting the cart before the horse. Consider the case of individual agency: it is not because you have a decision procedure and clear purpose that you are an agent. Rather, it is because you’re an agent that you (should) have a decision procedure and a clear purpose. Why not assume that the same is true in the plural case? Why assume that collections have to be well-organized and single-minded in order to be agents, while in the case of individual agency, we obviously assume that there is some “basic” sense, some “core” of agency in virtue of which some degree of organization and agenda can be *demand*ed of them (at least where morally sensitive issues are concerned)?

To answer this question, we obviously need to know what this mysterious “core,” “basis,” or “principle” of agency (that feature of which decision procedure and purpose, where they exist at all, are just symptoms) really is. Let us consider the individual case first. What exactly is it in virtue of which you blame an entirely disorganized individual with a scattered monkey mind for not “making up his or her mind” and getting his or her act together, rather than treating him or her as a being that is not of the kind of an agent because he or she has no established decision procedure or rational point of view? In order to see the decisive feature, and in order to understand how it might be at work even in some random collections, we obviously need to dig a little bit deeper into the structure of agency than most current accounts of group agency like to do. Virginia Held herself shows where to start. In her discussion of the general notion of responsibility, Held points out that responsibility implies some sort of practical knowledge, knowledge of what it is the agent is doing:

To hold an individual responsible for an act requires that he be aware of the nature of the action, in the sense that he is not doing A in the belief that he is doing B. (Held 1970, 472)

Of course, this cannot mean that you’re off the hook as long as you did not think what you were doing was morally reprehensible. All it does is to preclude holding an agent responsible for what Held calls “thoroughly unascertainable”

aspects of your action. To use Held's example: if you throw an explosive through the window of a house, you're morally responsible for the death of the inhabitant, even if you didn't know that somebody was home and if you performed your action simply in the belief that you're blowing up a house, not in the belief that you're killing the inhabitants. In this sense, your belief about what you're doing does not limit your responsibility, and it is easy to see why: if you didn't know you were killing the inhabitants, you *should* have known, and this is why you're still responsible. But if you ring a doorbell that unbeknownst to you has been connected to an explosive in the house by some villain, you're not morally responsible for the death of the inhabitants. You still killed the inhabitants, but not only did you not know, but it is also wrong to say you *should* have known—and that's why you're morally off the hook in this case. Held stops her discussion there, and she proceeds to examine how this translates to the case of a random collection. But it is well worth dwelling on the normative nature of the epistemic condition of responsibility a bit further. What is the feature in virtue of which it makes sense of an agent that in some cases, he or she *should* have known what he or she was doing, even if he or she did not in fact know? In the case of our example, an obvious answer seems to be that what the agent *did* know was that using explosives is dangerous, that houses often have people in them, that putting other people's lives at risk is wrong, and so on. Extrapolating from such intuitions, it seems plausible to say that if we hold people responsible on the basis of what they *should* have known rather than on the basis of what they actually knew, we ultimately do so in virtue of things they *did* know, and from which they *should* have drawn (but culpably *failed* to draw) the right conclusions. It is in this sense that actual "awareness of the moral nature of the action" is indeed an epistemic condition of responsibility.

How, then, does this translate to the plural case? Arguing for the view that random collections of individuals can be morally responsible, Held argues such collections can indeed be said to have "awareness of the moral nature of the action." This is an interesting suggestion well worth exploring. Yet Held brings up the issue only to add the following rather hasty remark:

If we say that, in special circumstances, a random collection can be aware of the moral nature of an act, we do not claim the existence of an inexplicable group awareness over and above the awareness of its individual members, only that we are sometimes entitled to say 'Random collection F is aware that p,' even though we cannot carry out a reductionist demand for statements about each individual member. (Held 1970, 476)

The context of this passage suggests that Held thinks the nonreducibility of such statements is due to pragmatic reasons only: we just can't say who's actually individually aware and who's not, that's all. No special sort of "group awareness" is needed, just some distribution of individual awareness—a distribution in which individual shares may be difficult to ascertain.

It is at this point that Held misses the crucial point concerning the kind of collective responsibility that can apply to random collections of individuals, or so I shall argue in the following. To see the point, let us once again return to the singular version of the knowledge or awareness in question. Remember that the knowledge or awareness in question is the feature in virtue of which an agent is responsible for an action. As we have argued, this feature must also be the feature in virtue of which an action is a subject's *own* action (in the sense in which it is only for their own actions that we blame agents). The "ownness" of the action in question has to be part of what is known—or of which there is awareness—by the agent, in the knowledge or awareness in question. In other words, this knowledge or awareness (or consciousness, in Held's sense) has to be *self*-knowledge, *self*-awareness, or *self*-consciousness, and, as we shall see shortly, it has to be self-knowledge, self-awareness, or self-consciousness *of a special kind*.

At this point, an important and well-analyzed distinction is crucial. Knowledge, awareness, or consciousness of ourselves can be *de re* or *de se*, observational versus nonobservational, reflective versus pre-reflective, first-personal versus third-personal (I take all of these terms and distinctions to be equivalent in the following). The question of which of the two *kinds* of knowledge often matters (as it does in our case, as I shall argue). The *locus classicus* in the received literature for a case of self-knowledge of the *de re*, observational, reflective, or third-personal kind is Ernst Mach (1886, 3) who enters a bus and sees his own image reflected in a mirror, and without knowing that it is himself he observes, judges that his clothing is shabby. Looking at the mirror, he learns something about himself, but the knowledge in question is not of the first-personal kind, and it is thus deficient. Mach thinks that his case shows how knowledge and recognition work first and foremost through general schemes and stereotypes rather through recognition of particular facts and specific details. Though this is certainly an important point to make, Mach thereby missed the crucial point of his own case. Perhaps it would have *led* to a recognition of himself as himself had he focused on the details of his own appearance rather than on the general "shabby schoolmaster" type. But no, however, specific detail would have *constituted* that knowledge. Though it is *unlikely*, it is *possible* even to read one's own name on a tag in a mirror, and still fail to know that the person is oneself.

Is the same true for practical knowledge? Is it relevant for the question of responsibility? Consider the following moral equivalent of Ernst Mach's case. Imagine a version of the above scenario in which Mach, again, enters the bus and sees himself mirrored in the opposite window without recognizing that it is himself that he sees. Now assume he observes the following: that man in the mirror is blocking an elderly person's way; the elderly person struggles with climbing the stairs. Mach sees what is going on and correctly judges that the man in the mirror is doing something wrong; he should be more circumspect, he should realize what he is doing to the people behind him by just standing there, and he should move on. Mach has thus clear knowledge and sound judgment of the moral nature of what he is doing, but he fails to know it *in the right way*: he

fails to *self-know* it, he does not know that *it is he* who is doing the blocking. Even though there is knowledge of the moral nature of his act *of sorts*, this kind of knowledge does not, in itself, constitute moral responsibility. Certainly, we would say that the Mach of our example *should* have been more circumspect, and perhaps that he should have realized that it was *of himself* that he realized that he blocked the elderly person's way. But if we assume of Mach that in the given situation, he was nonculpably "self-blind" (Shoemaker 1996, 30), or that the mirror image was such that we could not reasonably expect Mach to recognize his mirror image, we have to exempt Mach from moral responsibility for blocking the elderly person's way, even though it remains true that he was fully aware of the moral nature of what he was doing. The example illustrates that for somebody to be morally responsible for what she does or fails to do, it is not enough for that agent to know the moral nature of what she is doing or not doing *de re*, *observationally*, or *third-personally*; she needs to know the moral nature of what *she* is doing or not doing *de se*, *first-personally*, or *non-observationally*.

To see how this translates to the plural, consider first the following variation of Held's subway car example. We are a group of passengers in the subway car, and we're looking out of the window. There is another subway train running at the same speed on the neighboring track. The other subway train is unlit, but as the lights of our own train are mirrored in the other train's windows, we're seeing what's going on in our own train, thinking it's happening in the other train. Assume that what we're seeing is the bully abusing his victim—and we are outraged at the fellow passengers whom we see looking attentively out of the window instead of coming to the victim's aid. In this case, we are fully "aware of the moral nature of the action," and we strongly condemn what we are doing; but we fail to be aware of the action—or rather, omission—as *ours*: the knowledge or awareness in question is not of the first-personal kind.

What exactly is that plural first-personal awareness that is missing in our parallel train case? One might think it is just like the morally problematic version of Mach's bus case mentioned above, only that it also involves some other people. Just like our Mach, looking in the mirror, fails to know that is *him* who is blocking the elderly person's way, it seems that in the parallel train case, each of us fails to know that it is his or her *own* train he or she is looking at. In this view, the way out of the misconception is for at least one of us to recognize *him- or herself* in the mirror image, and then to alert the others. But this does not seem necessary. In this case, you need not recognize *yourself* in the mirror. Recognizing any other fellow traveller, or just our own train, will do. The required reasoning need not run through a moment of "oh, it's me!," as there are other equally plausible ways to the decisive insight: "Oh, it's *us*!"

III

It is often tacitly assumed in the literature that first-personal practical knowledge exists only in the singular. This creates difficulties for the analysis of

joint action. If practical knowledge of what it is *we* are doing, or failing to do, is really just first-personal knowledge of what *I* am doing, or failing to do, together with third-personal knowledge of what other people are doing, it seems that the only kind of *own* action that exists is *mine*, which seems to condemn us to a view in which there is never truly such a thing as an action that is *our own*. All that remains to do for the analysis of joint action, on this view, is to identify complex structures of singular actions that *look like* they are truly ours.

What we see in the discussion on collective responsibility is really just another symptom of this more general tendency. And the worry that seems to prevent many from accepting plural self-knowledge is clearly expressed in Held's above statement. It seems that for *us* to be plurally self-aware of an action or omission *as ours*, there has to be something like a "we" over and above the heads of the individuals. In another paper (Schmid, 2017), I argue that this view is wrong, and comes from a general misconception of the subject of intention. Just as in the individual case, there is no such thing as "the I" somewhere in my head that has my intention, we need not assume that there is such a thing as "the we" when we intentionally act together. Being first-person plurally self-aware of an action or omission *as ours* is not a property of some preexisting "we"—rather, it is the core of plural agency itself. The central claim is thus this: In order to be collectively responsible for an action we, together, need to be aware of our action (or omission) *as ours*, collectively. This knowledge, awareness, or consciousness is first-personal, but it is first-personal *in the plural*. It is not a combination of each participant's self-awareness of what he or she is doing, individually, with some observational or inferential knowledge of what others are up to. I have argued in a series of papers that, while group self-knowledge, or plural self-awareness, certainly differs from individual self-knowledge, or singular self-awareness, it is of the same kind as the latter (Schmid 2014b; 2014ab; 2016). Self-awareness, self-consciousness, first-personal knowledge comes in two forms: singular and plural.

Held ignores this because first, she does not account for the first-personal nature of the awareness or knowledge she requires for responsibility in general, and second, because she thinks that if that kind of awareness were different from individual awareness or knowledge, it would have to be "an inexplicable group awareness over and above the awareness of its individual members." The mistake to be found here, as in so many other cases, is the idea that if a group has awareness or knowledge, it has to be a collective *singular* awareness or knowledge. Yet the plural is not a collective singular. Plural self-awareness is something individuals have, not somebody else over and above their heads, only that in contrast to singular self-awareness, they have it only *together*, as a group.

Let us now turn back to our subway car example to see how this spells out in terms of responsibility. Imagine we are that random collection of strangers. Take the asymmetrical scenario. You are the one member in the collection that does everything that is possible "to transform the collection into an organized group." I am one of those members who looks away and does not do anything. It

was argued above that the question of whether or not we are collectively responsible for the failure to help depends on whether this failure can be attributed to us, collectively, as *our* failure. And it was claimed that this depends on whether or not we, together, are plurally self-aware of the failure *as ours*. How do we *find out* whether or not this is the case? One condition of the plural awareness in question is joint attention. If you're aware of what is going on, and of the fact that assistance is needed, but you think that I'm unaware of it, your responsibility is to establish joint attention and readiness for action. If we're jointly aware of what's going on, and I refuse to participate, I am personally responsible for my refusal, and thus for our failure. Yet the failure itself is still *ours*. *We* failed, and we are thus the proximate target for blame. But in this scenario, you're personally off the hook, because you did not do anything wrong—it is me who did wrong, and the wrong is of a special kind: I prevented *us* from doing the right thing, and thus I wronged *you*, too. Thus I'm doubly guilty; I did not engage in mobilizing the group, and by not responding to your initiative, I prevented *us* from doing what is right, thereby wronging *you*, the willing cooperator. And yet, this special form of being personally responsible depends on a failure to act which is genuinely *ours*.

In a version of Held's initial scenario in which there is no joint attention, the structure of responsibility might be different. Each of us is personally responsible for not taking action to mobilize the group. But assuming joint attention, is also the case that we, together, are collectively responsible for failing, as a group.

To test this claim consider the following cases from the victim's perspective and judge for yourself whether or not you accept the consequences of my account. Assume the bully stops his abuse and leaves the train. You're left bruised and angry at your fellow passengers who failed to team up and come to your aid. Assume first that the following conditions hold. The subway car is designed in such a way that there are separate compartments, with no more than one passenger placed in each compartment. The passengers are seated in their respective compartments in such a way that each passenger could see the abuse, but no passenger could see any other passengers. In this scenario, Held's distributive analysis seems to work neatly—as the victim, you'll be angry at *each individual* for not looking for help and thus failing to realize the conditions of team action. Depending on how severe your bruises are, and depending on your character, you might be willing to forgive each of your passengers, especially if each of them honestly asks for it, and that might be all the forgiveness needed in the case, there is nothing more to forgive.

The situation looks different in a second scenario. Assume now a different subway car design, one that is such that it is all in the open (no separate compartments, two rows of seats alongside the sides of the car). In this case, not only could everybody see what is going on, but everybody could see that everybody could see it, and so on. In this scenario, you might not just be angry at *each of them* after the abuse has finally stopped and the bully has left. In

addition, you might be angry at *all of them, together*. What you need might not just be an apology from *each of them*, separately (though this will certainly be needed if nobody did anything to initiate a team action). In addition, you might want something more: an apology from *all of them, together*, some sort of them apologizing to you jointly and publicly rather than individually and in private conversation with you, because this is how they failed: not just individually (by failing to look for potential cooperators), but *collectively*, as a group (by failing to act). Since you did not receive any help from them, together, you might now want them to get their joint act together at least in asking jointly for your forgiveness.

Let me conclude with a brief remark on the connection between collective responsibility and group agency. It is obvious that a random collection of individuals is not a group agent of the received kind. There is no joint commitment, rationally unified perspective, or decision rule in this case. But even unorganized random collections can be plurally self-aware of what should be done as *that collection's own* actions of the sort that constitutes responsibility. This is the case wherever we, jointly, know first person plurally what we should or should not do together. As I have argued elsewhere, plural pre-reflective self-awareness, or groundless group self-knowledge, is not in itself the kind of organization required for the received group agents. But it is the feature in virtue of which there is normative pressure toward rational unification and organization among us. It is in virtue of our self-awareness that we are the owners of our attitudes and actions, in virtue of which there is pressure for formal and material unity of our attitudes, and in virtue of which our attitudes are our commitments (Schmid 2014a; 2014b; 2016). This has a singular or individual as well as a plural or collective form. And just as we are individually responsible for our actions, as agents, even if we fail to achieve full unity, we are collectively responsible even if we fail to organize ourselves. Plural agents do not exist in virtue of some achieved rational unification and institutionally established organization, but in virtue of the feature which provides the normative pressure toward rational unification and organization. In this sense, even random collections of individuals can, under some conditions (such as joint attention), be collective agents, and thus indeed be collectively responsible. And in this sense, understanding how random collections can be collectively responsible is important for understanding *organization*, and thus the collective responsibility of organized group agents. Organizing ourselves, collectively, and getting our act together as a team is something we can only do together, and it is thus, under suitable circumstances, our collective duty.

IV

This research was supported by Austrian Science Fund (Grant number: I 3068 – G24) [Correction added on 24 September 2018, after first online publication: Funding information has been added.].

Notes

- ¹Held construes the case a bit differently in that she has the second smallest of seven people in the subway car throwing the smallest passenger to the ground and slowly beating and strangling him or her to death—this particular ordering by size does not help to make the case particularly obvious as a case of duty to act jointly, especially given the gravity of the crime being committed—though she adds that in the scenario she has in mind, “no one of the five, acting alone, could have subdued him” (Held 1970, 477).
- ²In fact, this is how Held (1970, 471) defines “random collections”: “For this discussion, I shall mean by a ‘random collection of individuals’ a set of persons distinguishable by some characteristics from the set of all persons, but lacking a decision method for taking action that is distinguishable from such decision methods, if there are any, as are possessed by all persons.”

References

- Bates, Stanley. 1971. “The Responsibility of ‘Random Collections’.” *Ethics* 81: 343–49.
- Bostyn, Dries H., and Roets, Arne. 2016. “The Morality of Action: The Asymmetry between Judgments of Praise and Blame in the Action-Omission Effect.” *Journal of Experimental Social Psychology* 63: 19–25.
- Chant, Sara Rachel. 2015. “Collective Responsibility in a Hollywood Standoff.” *Thought* 4: 83–92.
- Gilbert, Margaret. 1989. *On Social Facts*. Princeton, NJ: Princeton University Press.
- Held, Virginia. 1970. “Can a Random Collection of Individuals Be Morally Responsible?” *Journal of Philosophy* 67: 471–81.
- List, Christian, and Pettit, Philip. 2011. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press.
- Mach, Ernst. 1886. *Analyse der Empfindungen*. Jena, Germany: Gustav Fischer Verlag.
- Rovane, Carol. 2004. “What Is an Agent?” *Synthese* 140: 181–98.
- Schmid, Hans Bernhard. 2014a. “Plural Self-Awareness.” *Phenomenology and the Cognitive Sciences* 13: 7–24.
- . 2014b. “Expressing Group Attitudes. On First Person Plural Authority.” *Erkenntnis* 79: 1685–1701.
- . 2016. “Groundless Group Self-Knowledge.” In *The Epistemic Life of Groups: Essays in the Epistemology of Collectives*, ed. Michael S. Brady and Miranda Fricker, 51–74. Oxford: Oxford University Press.
- . 2017. “The Subject of ‘We Intend’.” *Phenomenology and the Cognitive Sciences*. Advance online publication. <https://doi.org/10.1007/s11097-017-9501-7>
- Shoemaker, Sidney. 1996. *The First-Person Perspective and Other Essays*. Cambridge: Cambridge University Press.
- Zimmerman, Michael J. 1985. “Sharing Responsibility.” *American Philosophical Quarterly* 22: 115–22.