

The Ethics of Connected Minds
Jeff Sebo and Luke Roelofs

Abstract

Suppose that minds can be connected, in the sense of sharing direct mental connections or even token mental states. How does the idea of connected minds affect philosophical thinking about the self, welfare, rights, virtues, and other such matters? In this paper, we survey these questions and motivate a general proposal about how to answer them, based in part on views developed in the literatures on collective agency and personal identity. Our general proposal is that moral properties and relations track mental properties and relations. If two subjects are partly mentally connected in relevant respects, then they are also partly morally connected in relevant respects. As a result, in a world of connected minds, parts of individuals, individuals, and groups of individuals can all have selves, welfares, rights, virtues, and so on at the same time. Moreover, as parts and groups of individuals become more important in psychological explanations, they will also become more important in moral explanations, perhaps even replacing individuals as the primary units of moral analysis in many contexts.

1. Introduction

Suppose that minds can be connected. In the same way that matter can be networked with other matter, minds can be networked with other minds. And in the same way that matter can contain other matter or be contained by other matter, minds can contain other minds or be contained by other minds. Of course, we know that at least some minds are connected now, for instance in the case of some conjoined twins. But what if the scope of connected minds is more expansive than we appreciate? And what if, with the advent of new technologies (particularly artificial intelligences and artificial-biological interfaces), the scope of connected minds will be even more expansive in the future? How would that affect our thoughts, feelings, and relationships? How would it affect our understanding of the self and personal identity?

The possibility of connected minds raises ethical questions as well. If you and I have connected *mental* properties and relations, do we also have connected *moral* properties and relations? For example, to what extent do we share benefits and harms, and to what extent do we retain individual benefits and harms? To what extent do we share rights and virtues, and to what extent do we retain individual rights and virtues? And to what extent do we count as the same unit of moral concern from the standpoint of

justice, and to what extent do we still count as separate units of moral concern from this standpoint?

While some of these questions are familiar in the context of collective ethics and ethics and personal identity, not all of them are. And either way, they take on new urgency in a world of connected minds.

In this paper, we survey these questions and motivate a general proposal about how to answer them. Our general proposal is that moral properties and relations track mental properties and relations. If two humans are partly mentally connected and partly mentally separate in relevant respects, then they are also partly morally connected and partly morally separate in relevant respects. We already attribute duties, rights, virtues, and so on to groups of humans who are sufficiently connected, as well as to parts of humans who are sufficiently fragmented. We propose that we generalize these ideas, and that when we do, we can see that morality is much more complex, layered, and dynamic than many ethicists assume. In a world of connected and fragmented minds, connected and fragmented moral subjects are everywhere.

We start, in section 2, by surveying the metaphysics of connected minds, discussing how minds can be networked, how minds can be combined, and what might follow for the self and personal identity. Then, in section 3, we survey the ethics of connected minds, asking whether mental connections lead to moral connections, and what might follow for morality if they can. Then, in section 4, we sketch and motivate our general proposal that (relevant) mental connections do, in fact, lead to moral connections. To be clear, our goal in this paper is not to fully establish this thesis. That would take much more space than we have. Our goal is instead to raise questions about the ethics of connected minds, show why these questions are important, and suggest one possible answer, with the hope of generating further discussion.

2. The Metaphysics of Connected Minds

Many kinds of mental connection might be possible. We will focus on two kinds here for the sake of simplicity. First, *networked minds* are minds which do not overlap but do stand in a ‘direct’ mental relation. To take a simple example, if you and I can communicate via telepathy, then we have networked minds, insofar as my thoughts are able to ‘directly’ make contact with yours, with no need for

intermediate actions or perceptions. Second, *combined minds* are minds that overlap with each other, in the sense that they share token mental states. To take a simple example, if the left and right hemispheres of your brain turn out to be separate minds *as well as* parts of one and the same mind, then you have combined minds: the mental states of the whole brain are shared with its two halves.

Some of these connected minds might exist now. Take octopuses. Octopuses have one “central” brain as well as a “peripheral” brain in each arm, and they exhibit semi-integrated and semi-independent behavior. If each brain is capable of independent thought and feeling as well as direct communication with the others, then these brains count as networked minds in our sense. And if all this brain activity contributes to a general set of thoughts and feelings experienced by the octopus as a whole, then these brains also count as combined minds in our sense of the term. As we will see, there are many other examples as well. On some theories of consciousness, humans, animals, and even plants might count as having networked or combined minds, whether or not any particular mind appreciates that fact.

And whether or not connected minds exist now, they might exist in the future. Take artificial intelligences. Many computer programs are networked with each other and layered on top of each other. If and when these programs become conscious, it might be that they develop a shared set of thoughts and feelings while retaining separate sets of thoughts and feelings. Additionally, we already have technology that can link human brains with computer programs. As this technology matures, we might discover that we can connect biological and artificial minds in these ways, or even that we can connect biological minds with other biological minds, via connections with computer programs. Whether or not these developments are advisable, the ethical questions that they raise are worth exploring.

This section surveys the kinds of networked and combined minds that might exist now or in the future, as well as the questions that these minds raise about the self and personal identity. To be clear, this survey will not be exhaustive. There might be important kinds of mental connection that it omits. But it will be enough to establish that connected minds can come in many different shapes and sizes, and that they might increasingly be the rule, not the exception. So, philosophical questions about connected minds

might matter not only in cases involving, say, dissociative identity disorder and teletransportation, but also in cases involving neurotypical humans and nonhumans. And even if these questions only matter for some current beings, they might matter for many future beings, human and nonhuman alike.

2.1. Networked minds

We can start with networked minds, which are minds that share a direct mental relation. A paradigmatic example is telepathy. If someone can interact with your mental states directly, via some entirely mental sequence of causes and effects, rather than indirectly, via, say, spoken language, then you have networked minds in our sense of the term. Of course, if two minds are sufficiently networked in this way, then the question arises whether they still count as distinct minds at all, as opposed to (merely) distinct parts of a single mind. We will consider this question below. But this rough gloss will be enough for now.

Intuitively, two lakes can share water directly, via a tributary, while still counting as two lakes. Similarly, two minds can share mental states directly, via mental transfer, while still counting as two minds.

Of course, directness and indirectness come in degrees. And there is a lot of variation in this regard, in both characteristically inter-mind relations and characteristically intra-mind relations. On the inter-mind side, we sometimes need to take relatively many steps to convey mental states to someone else, such as when two people who speak different languages need to communicate even simple ideas via an interpreter. But other times, we can convey mental states with fewer intermediaries, such as when a married couple can communicate complex ideas via a raised eyebrow. The intuitive draw of the idea of a “group mind” may reflect the awareness that when people know each other well enough, we can cut out many of the normal steps of interpersonal communication and share ideas relatively directly.

Conversely, on the intra-mind side, we can sometimes take convey mental states to ourselves in relatively few steps, such as when we store information in short term memory such that we can access it at will. But other times, we need to take relatively many steps, such as when we store information in our

phones such that we can access it only by entering our password, opening the folder, and perceiving the information. Indeed, a common complaint about some new technologies is that they create a state of dependence, such that we increasingly “outsource” tasks like remembering phone numbers, and so we increasingly need to coordinate with our past and future selves in relatively indirect, characteristically interpersonal ways rather than in relatively direct, characteristically intrapersonal ways.

Given this variation, it can be surprisingly difficult to draw a clear line between networked minds and non-networked minds in our sense of the term. But for our purposes here, it will be enough to say that if two minds are capable of communicating via at least some characteristically intra-mind relations, such as storing information in memory, rather than *merely* via characteristically inter-mind relations, such as storing information in a phone, making facial expressions, or speaking via an interpreter, then these minds count as networked in our sense of the term. Of course, more analysis is needed to sharpen this criterion, as well as to determine whether we should be attempting to draw a clear line between networked and non-networked minds in the first place. But a rough criterion is all we need for now.

Of course, the idea of networked minds is familiar from the literature on personal identity. On some views about personal identity, if a person divides into two people, such that each successor inherits all the beliefs, desires, intentions, memories, and so on of the predecessor, then (a) the successors count as having separate minds from the predecessor, but (b) the successors also have a strong diachronic psychological connection to the predecessor, not only in virtue of having similar mental states but also in virtue of having acquired these mental states via a characteristically intra-mind sort of interitance. Our present suggestion is that we can stand in such relations not only in cases involving, say, fission and fusion but also in synchronic cases involving other kinds of links between minds.

At present, the clearest example of networked minds in humans is the Hogan twins, “whose brains are connected by a thalamic bridge” that allows them to share mental states. For instance, they can share perceptions: “Tatiana can see out of both of Krista’s eyes, while Krista can only see out of one of Tatiana’s.” They can share thoughts: “The twins say they know one another’s thoughts without having to

speak. And they can share motor control: “Tatiana controls three arms and a leg, while Krista controls three legs and an arm. They can also switch to self-control of their limbs.” Assuming that these self-reports are reliable, each twin has a mind of her own, but they can also communicate via characteristically intra-mind relations. This is enough for them to count as having networked minds in our sense.

A possible example of networked minds in nonhumans involves distributed cognition in invertebrates. To return to an example mentioned above, octopuses have a relatively large “central” brain as well as a relatively small “peripheral” brain in each arm, and they behave in a way that suggests that these brains are partly integrated and partly fragmented. While we can only speculate about what, if anything, it might be like to be an octopus, one possibility is that each of their nine brains constitutes a mind of its own, such that the peripheral minds can relay sensory information back to the central mind, which can then relay motor instructions back to the peripheral minds. In this scenario, octopuses would count as having multiple networked minds rather than as (merely) having a single spatially distributed mind.

And of course, new technologies might allow for many more networked minds in the future. Humans are currently developing advanced AIs, and if and when advanced AIs become conscious, then they will be capable of having networked minds in all the same ways as human and nonhuman animals, as well as, perhaps, other ways. Humans are also developing technologies that aim to approximate telepathy, such as the secretive US military “Silent Talk” project. If and when these technologies succeed, then new kinds of networked minds might be possible both within and beyond humanity. In some possible futures, such as futures where humans (or our successors) find it useful to scale up networked minds or telepathy technologies, these cases of networked minds will increasingly be the rule, not the exception.

2.2. Combined minds

We can now consider combined minds, which are minds that share a direct *compositional* mental connection. That is, they contain each other, overlap with each other, or otherwise share mental parts – that is, token mental states – with each other. A clear example would be: Suppose that your left and right

hemispheres are both *individually* and *jointly* conscious. That is, each hemisphere has a unified consciousness, and these consciousnesses constitute the unified consciousness of your brain as a whole. In this case, you would count as having combined minds in our sense. As before, this idea raises difficult metaphysical questions, which we will consider below. But this rough gloss will be enough for now. We all agree that objects can overlap in such ways. Our suggestion is that minds can do so as well.

As with direct causal relationships, direct compositional relationships can come in degrees. For instance, when minds stand in part-to-whole relations with each other, the parts can make up less or more of the whole. Suppose that the “part-minds” each constitute only one percent of the mental states of the “whole-mind,” kind of like the citizens in a small town participating in the democratic process. In this case each part-mind would play only a minor role in determining what the whole-mind is like. In contrast, suppose that, as in the case above, the part-minds each constitute fifty percent of the mental states of the whole mind, kind of like a married couple (without children) making a family decision together. In this case the part-minds would play a major role in determining what the whole-mind is like.

Similar remarks apply to minds that overlap, such that they have some shared mental states while each having some separate mental states. In particular, when only a low percentage of mind A belongs to mind B and vice versa, such that the overlap is barely there at all, these minds are only weakly combined. When a high percentage of mind A belongs to mind B and vice versa, such that the overlap is nearly complete, these minds are strongly combined. And when a low percentage of mind A belongs to mind B but a high percentage of mind B belongs to mind A, these minds are weakly combined in one respect (insofar as the question is how much mind A constitutes mind B) but strongly connected in another respect (insofar as the question is how much mind B constitutes mind A).

As some of these examples suggest, the idea of combined minds is somewhat familiar from the literature on collective agency. On some views about collective agency, if two people share intentions in the right kind of way (say, by each intending to participate in a shared activity), then (a) these agents count as distinct agents, with distinct intentions, and (b) these agents count as a collective agent (sometimes called a ‘plural subject’), with a shared intention, such that the former intentions jointly

compose the latter intention. Part of our suggestion is that, at least in principle, this kind of relation can arise not only for functional states such as intentions, but also for phenomenal states such as experiences. Two beings can have distinct *and* shared experiences, in different ways and to different degrees.

A possible example of combined minds in humans involves dissociation. We all have sets of mental states that are associated with each other, such that when some are activated, others are too. In many cases, these sets of mental states overlap enough that we find it natural to classify them as parts of the same mind rather than as different minds. But in the case of humans with some dissociative disorders, these sets of mental states become compartmentalized enough that some of us find it natural to classify them as different minds. If we suppose that at least some humans have sets of mental states that are (a) compartmentalized enough to count as different minds yet (b) connected enough to share at least some token mental states, then they count as having combined minds in our sense of the term.

At least in principle, this phenomenon can arise in non-pathological cases as well. Suppose that you have overlapping minds: a “work self” that activates at work, a “home self” that activates at home, and so on. These minds overlap enough that they experience themselves as parts of a single, unified mind, and they use the first-person singular to refer to this single, unified mind, for instance by using terms like “me” to refer to the person as a whole and terms like “my work self” and “my home self” to refer to the parts of that person. In this case, you might interpret yourself as having a single, unified mind. But this interpretation might be, as Dennett puts it, a useful fiction. The reality is instead that you have multiple overlapping minds and you construct a single, unified self-narrative to make sense of them together.

Many other examples are possible too, depending on our views about cognition, computation, and consciousness. On some views about these matters, parts of organisms can be conscious, organisms can be conscious, *and* groups of organisms can be conscious. For instance, we might think that octopuses have more minds than brains, that is, that each octopus brain has a mind, and that the octopus as a whole has a mind as well. We might also think that insect colonies have more minds than brains: Each insect has a mind, and the colony as a whole has a mind as well. If plants are conscious, perhaps the same can be

true for them: Each plant has a mind, and a network of plants has a mind as well. And perhaps in some cases, all three can be true at once: parts, wholes, and groups can *all* have minds.

As before, new technologies might allow for many more combined minds in the future. In addition to creating AIs who can communicate with each other via characteristically intra-mind relations, we can also create AIs who can contain other AIs, overlap with other AIs, or otherwise share token mental states with other AIs. Perhaps new technologies will allow humans to relate to each other, other animals, and AIs in this way too. Already, humans who store information on a shared Google document have at least a weakly combined mind, in that we share at least some token mental states, according to the extended mind hypothesis. If and when Google Drive updates to allow for stored conscious experiences, these mental combinations could extend according to this hypothesis.

1. Metaphysical issues

As noted above, this rough description of connected minds raises a wide range of metaphysical questions that extend beyond the scope of this paper. One set of questions, of course, is about the status of particular beings. For instance, *if* conjoined twins, humans with multiple personalities, and other such beings function roughly as described above, *then* they count as having networked or combined minds in our sense of the term. But further discussion is needed to establish that they do, in fact, function this way. More importantly, if our focus is on conscious minds in particular, then how many connected minds there can be depends in part on which theory of consciousness is true. And the metaphysical status of these minds depends in part on which theories of the self and personal identity are true.

Consider first the relevance of theories of consciousness. On one hand, if we accept a narrow theory of consciousness, such as identity theory, then we might think that some physical systems, such as animals, can have conscious minds, but that other physical systems, such as plants, cannot. In this case, we might still think that connected minds can exist, for instance in cases where parts of animal brains can be conscious or in cases where animal brains can be connected either naturally (as with the Hogan twins

or octopus brains) or artificially (via new technologies). But the scope of combined minds might be relatively small. For instance, we might think that plants, AIs, and many parts and groups of animals (to say nothing of parts and groups of plants, AIs, and other beings) are all ruled out.

On the other hand, if we accept a wide theory of consciousness, such as panpsychism, then we might think that *all* physical systems count as conscious. In particular, we might think that connected minds can be as common as connected matter: To the extent that objects can be networked or combined, minds can be networked or combined as well. Thus, not only can we identify minds in parts of organisms, whole organisms, and groups of organisms, but we can also identify minds in all other parts, wholes, and groups. In this case, we might still think that some minds are more important than others for purposes of explaining, predicting, and controlling particular phenomena. But the metaphysics of the mind would be every bit as complex, dynamic, and layered as the metaphysics of matter.

Of course, there are many options between these extremes as well. For instance, standard functionalist theories plausibly have wider implications than identity theory and narrower implications than panpsychism. In particular, these views plausibly imply that *some but not all* parts, wholes, and groups count as having conscious minds. But even if many current parts, wholes, and groups fail this test, many future parts, wholes, and groups will plausibly pass it. The spectrum of possible minds is much wider than the spectrum of actual minds, and new technologies have the potential to realize new possibilities. So on this view, we might think that we have a moderate number of connected minds now, and that we have the potential to create a much higher number of connected minds in the future.

Second, to the extent that connected minds exist, they raise difficult questions about the self and personal identity. When minds are connected causally or compositionally, what is the metaphysical status of these minds? Do they count as separate minds, do they count as parts of the same mind, or do they count as both at the same time? And insofar as minds count as having or being selves and persons, what is the metaphysical status of these corresponding selves and persons? When minds are connected causally or compositionally, do they count as having or being separate selves or persons, do they count as having or

being parts of the same self or person, or do they count as both at the same time? While these questions extend beyond the scope of this paper, we can make a few remarks about them here.

First, some theories of the self and personal identity are compatible with the idea of connected selves and persons. Take the idea of the self as a center of narrative gravity. On this view, the self is a useful fiction, a simple abstract object that we posit for purposes of explaining, predicting, and controlling complex phenomena. And in the same kind of way that we can have multiple, overlapping centers of gravity, we can also have multiple, overlapping centers of narrative gravity. That is, I might find it useful to explain my behavior by thinking about parts in some cases (“this is what my work self prefers”), the whole in other cases (“this is what I prefer”), and groups in other cases (“this is what my family prefers”). Insofar as I do, I might count as having, or being part of, multiple, overlapping selves in this sense.

Second, however, other theories of the self and personal identity are not compatible with the idea of connected selves or persons. For example, a common class of views about personal identity holds that two beings are the same person, or as parts of the same person, if they have sufficiently strong bodily, psychologically, or narratively connectedness or continuity with each other in the right kind of way. And if two beings are the same person, or parts of the same person, then they are not also separate persons. On this class of views, when particular kinds of (causal or compositional) mental connections across minds are strong enough, these minds count as the same person, or parts of the same person. Otherwise they count as separate persons who happen to have an especially intimate relationship.

Since we will not be able to discuss these issues further here (though we each discuss them elsewhere), we will simply make two general points. First, we will assume a standard functionalist view about consciousness in what follows, though nothing of substance will turn on this assumption. Our analysis will hold no matter what; the only question is how extensive its implications will be. Second, we agree with Parfit and others that we should ask metaphysical and normative questions separately. So *if* we interpret selves and persons as metaphysical rather than normative entities, then we should keep an open

mind about whether selves and persons are the relevant units of moral analysis for particular purposes.

And our primary interest in this paper is in the latter question.

3. The Ethics of Connected Minds

Having surveyed the general idea of connected minds and some of the metaphysical questions that connected minds raise, we can now survey some of the ethical questions that connected minds raise. When minds are connected, do they have connected welfares, and do they have different amounts of welfare? Do they have connected rights, and do they have different rights? Do they have connected virtues, and do they have different virtues? And underlying all these questions are foundational questions about moral individuation. What are the basic units of moral analysis? And when different units of moral analysis are connected with each other, do they remain separate units of moral analysis, do they become (parts of) the same units of moral analysis, or both?

As the above discussion suggests, some of these questions are familiar from the literatures on collective agency and responsibility and ethics and personal identity. The former literature examines what happens when agents think and act collectively, and on standard views, it holds that different units of moral concern can be separate and connected at the same time; for instance, two people can have different individual rights and responsibilities as well as shared collective rights and responsibilities. The latter literature examines what happens when agents share bodily, psychological, or narrative continuity, and on standard views, it holds that different units of moral analysis become (parts of) one and the same unit of moral analysis once these bodily, psychological, or narrative relations become strong enough.

But some of these questions are also new or, at least, in need of modification. For example, the literature on collective agency and responsibility tends to presume that individual agents are conscious but that shared agents are not. How might our analyses change, or expand, if we presume that both individual and shared agents are conscious in some cases? Similarly, the literature on ethics and personal identity tends to presume that these questions arise primarily in cases involving split brains, multiple personalities,

psychological fission, psychological fusion, or substantial change over time. How might our analyses change, or expand, if we presume that these questions arise not only in these relatively exceptional cases, but also in the ordinary operations of parts, wholes, and groups of biological or artificial systems?

This section starts by surveying some of the questions about welfare, rights, and virtues that connected minds raise, along with some possible answers. As with the previous section, this survey will not be exhaustive. But it will be enough to illustrate the strange moral future that we might face. This section then closes by motivating, without fully establishing, our general view about the matter: When minds are partly connected and partly disconnected, moral subjects can be partly connected and partly disconnected too. And when parts of individuals, individuals, and groups of individuals all have minds at the same time, they can all be moral subjects at the same time. A world of networked and combined minds is thus also a world of networked and combined moral subjects.

1. Welfare

Take questions about welfare first. One question concerns individuation. When two minds are connected, are their welfares connected as well? That is, when one is benefited or harmed, is the other as well? Plausibly, it depends. If two minds share a token pleasure or pain, then (on hedonic views) it makes sense to say that both are benefited or harmed via that token state. Similarly, if two minds share a token desire or preference, then (on desire or preference satisfaction views) it makes sense to say that both are benefited or harmed via the satisfaction or frustration of that state. In contrast, if two minds are connected in general but not with respect to these welfare states in particular, then it makes sense to say that each mind can have its own benefits and harms, at least with respect to its own welfare states.

Of course, even if each mind has its own token welfare states, mental connections can still affect these welfare states. If two people care about each other, each will be happy to see the other happy and sad to see the other sad. Each will also be satisfied to see the other satisfied and frustrated to see the other frustrated, and insofar as they have similar goals, both will take satisfaction and frustration in similar

outcomes. If two minds are connected, these relations can all still hold, in many cases to a much greater extent. That is, in the same way that a married couple can have separate but intertwined welfares much more than strangers can, connected minds can have separate but intertwined welfares much more than unconnected minds can, in addition to potentially sharing token welfare states.

Another question concerns amounts. When two minds are connected, do they have different amounts of welfare? This question is partly descriptive. For instance, when two minds are connected in particular ways, can they experience more pleasure and pain, less pleasure and pain, or the same amount of pleasure and pain overall than they otherwise could? We can imagine a case for each answer. For instance, mental connections could have an amplifying effect for some welfare states, but could also consume mental energy that could otherwise be spent on welfare states. Either way, this question will matter a lot for utilitarians and other ethicists who care about aggregate welfare, since it will partly determine which kinds of future population can experience greater amounts of aggregate welfare.

This question is also partly normative. When two minds share a token welfare state, such that they both count as benefited or harmed via that state, should we “single count” or “double count” this state in aggregate estimates? For example, if two minds share one unit of pleasure, should we say that these minds have one or two units of pleasure in the aggregate? If we single count shared welfare states, then whether a world of connected minds has more welfare overall will turn entirely on the kinds of descriptive questions discussed a moment ago. Whereas if we double count, then a world of connected minds is more likely to have more welfare overall. Our own tentative view is that single counting makes more sense than double counting, but a lot depends on which theory of welfare one accepts and why.

2. Rights

Now take questions about rights. One set of questions, which correspond to questions about individuation, concerns how connected minds should relate to each other. Do connected minds have self-regarding

duties to each other, such as a duty to pursue each other's moral perfection? Or do they instead have other-regarding duties to each other, such as a duty to respect each other's autonomy? For instance, suppose that your day self and night self disagree about what to do: Your day self wants you to stay in, and your night self wants you to go out. Is your day self morally permitted to use self-regarding strategies for interacting with your night self, such as coercion? Or is your day self instead morally required to use other-regarding strategies for interacting with your night self, such as compromise?

Our own view, which we can only sketch here but one of us develops in more detail elsewhere, is that insofar as your personalities share intentions, they share self-regarding duties to themselves as a shared agent. But insofar as they have separate intentions, they still have other-regarding duties to each other as separate agents. Given that your personalities are psychologically connected, they can be expected to share intentions, and thus to share self-regarding duties, in a wide range of cases. But given that your personalities are less than fully psychologically connected, they can also be expected to have separate intentions, and thus to have other-regarding duties to each other, in a wide range of cases too. In this respect, they are no different in kind from persons who share agency, only different in degree.

A related set of questions, which again correspond to questions about individuation, concerns how connected minds relate to others. Do connected minds share duties to other moral patients, and do they share rights against other moral agents? For example, suppose your night self cheats on your partner. Is your day self morally responsible for this decision? This question is surprisingly hard to answer. On one hand, according to the fine-grained view about individuation that we sketched a moment ago, your personalities share duties only insofar as they share agency. So, if they jointly commit to cheating on your partner, then they share responsibility for this action. But if one of them cheats on your partner without the other's knowledge or consent, then only one of them is responsible for this action.

On the other hand, even according to this fine-grained view, we can still have reason to hold connected minds responsible for each other's behavior in some cases. There are several reasons why, but we can focus on one here to illustrate the point. Practically speaking, it makes sense for persons to interact

as persons, in the same kind of way that it makes sense for, say, nations to interact as nations. So, for instance, your day self can have a duty to apologize, on behalf of the person, for an action that your night self performed, in the same kind of way that a President can have a duty to apologize, on behalf of the nation, for an action that a past President performed. In both cases, whether or not this individual is fully *blameworthy* for the relevant action, they can still be *accountable* for it in virtue of their role.

3. Virtues

Now take questions about virtues. To what extent do connected minds have connected virtues and vices, and to what extent do they retain separate virtues and vices? If we understand virtues and vices as character traits, then this normative question reduces to a descriptive question: To what extent do connected minds have connected character traits? For instance, suppose that your day self and night self share a general tendency towards honesty, but that your day self tends to be more honest than your night self overall. In this kind of case, does it make sense to attribute a general character trait to both personalities (say, 80% honesty)? Or does it instead make sense to attribute a specific character trait to each personality (say, 85% honesty for your day self and 75% honesty for your night self)?

The answer to this question depends on which theory of character we select. But as above, we can sketch one possible answer for purposes of illustration. If we take the self to be a center of psychological gravity, then parts, wholes, and groups can all have selves. To see what we mean, consider centers of gravity. A center of gravity is the point in space around which an object's weight is evenly distributed. Tracking this point in space can be useful for purposes of explaining, predicting, and controlling an object's behavior in everyday life. And while we might find it useful to attribute centers of gravity to objects by default, we might also find it useful to attribute them to parts or groups in some cases, when the parts are sufficiently distinct or the groups are sufficiently connected.

Similarly, we can define a self as the simple, abstract set of mental states around which a mind's complex, actual mental states are evenly distributed. For instance, if your credence in a proposition fluctuates around .5, then you might think of yourself as having a credence of .5. As with centers of gravity, tracking this simple, abstract set of mental states can be useful for purposes of explaining, predicting, and controlling our behavior in everyday life. And while we might find it useful to attribute selves, in this sense of the term, to persons by default, we might also find it useful to attribute them to parts or groups in some cases, when the parts are sufficiently distinct (e.g., your day self and night self) or the groups are sufficiently connected (e.g., you and your partner).

This view allows for nuanced, layered moral assessments in cases involving connected minds. For instance, return to the case where your night self cheats on your partner. Does this action reflect badly on your day self too? Yes and no. If your partner accepts this view about character, they might respond as follows: "I know that you become someone else when you drink, and it was *that* version of you who cheated. But this version of you still *enabled* that version to cheat, and since these versions of you have a lot in common, what each does still reflects on the other at least somewhat. More importantly, I need to be partnered with *all* of you, not only *part* of you. And while I know that this version of you might not have cheated, the fact remains that your infidelity still reflects badly on you overall."

4. Units of moral analysis

In these brief discussions of welfare, rights, and virtues, we surveyed some of the questions that connected minds raise for ethics, as well as how we might start to answer these questions on a relatively fine-grained view about moral individuation. We now want to close by making a few related suggestions. First, fine-grained views about moral individuation have a general advantage over course-grained views, particularly in a world of connected minds. Second, we can dissolve some of the apparent implausibility of fine-grained views by making a distinction between *basic* and *default* units of moral analysis. Third, in

a world of connected minds, the default units of moral analysis might change, at least to a degree, and moral discourse and practice might need to change, at least to a degree, as a result.

As noted above, we can make a distinction between *course-grained* views about moral individuation, which set a low bar for two beings to count as (parts of) the same basic unit of moral analysis (that is, the same basic subject of welfare, rights, virtues, and so on), and *fine-grained* views, which set a high bar for two beings to count as (parts of) the same basic unit of moral analysis in this sense. For instance, a weak psychological connectedness view, which holds that any psychological connection is sufficient to make two beings (parts of) the same basic unit of moral analysis, is course-grained. In contrast, a strong psychological connectedness view, which holds that any psychological *disconnection* is sufficient to make two beings *different* units of moral analysis, is fine-grained.

In general, the benefit of course-grained views is that they capture much of our moral discourse and practice in a straightforward manner. That is, these views generally imply that persons are basic units of moral analysis (with the exception of cases involving fission, fusion, and other such disruptions), and so they generally vindicate our practice of attributing welfare, rights, virtues, and so on to persons. However, the cost of course-grained views is that they fail to make moral distinctions that, on reflection, seem warranted in some cases. For instance, in cases where a person's personalities think, feel, and act at cross purposes, fine-grained views are able to make distinctions between these personalities' welfares, rights, and virtues in a straightforward manner, whereas course-grained views are not.

With that in mind, we think that fine-grained views have a general advantage over course-grained views, which is that they can still vindicate the idea that individuals and groups can be proper units of moral analysis, even if they deny that individuals or groups are *basic* units of moral analysis. As noted above, philosophers standardly accept that when individuals think and act together, the group as a whole becomes a shared agent. According to fine-grained views, we can say the same about parts of individuals. For instance, we can say that when different personalities think and act together, the person as a whole

becomes a shared agent. So, we can still attribute duties, rights, and so on to individuals and groups. The only difference is that we should allow for the possibility that individuals are, themselves, groups.

Granted, fine-grained views might still seem intuitively implausible, since we standardly experience individuals as the basic units of moral analysis. But fine-grained views can explain away this implausibility, in at least two ways. First, as we have seen, fine-grained views are more plausible than course-grained views in some cases, such as when we find it natural to *both* attribute duties and virtues to particular personalities *and* attribute duties and virtues to the person as a whole. Second, even if fine-grained views are correct, we can still expect them to be implausible. Since we mostly interact with each other as individuals, it makes sense that we would experience individuals as the “primary” units of moral analysis, even if parts of individuals can be units of moral analysis too.

We can underscore this point by noting that the idea of “primary” units of moral analysis is ambiguous between two ideas that are worth distinguishing: *basic* and *default* units of moral analysis. On this distinction, a *basic* unit of moral analysis is a unit of moral analysis that is not, itself, composed out of other units of moral analysis. In contrast, a *default* unit of moral analysis is a unit of moral analysis that tends to be the relevant one for a particular agent in a particular context. These concepts are easy to conflate, but they can easily come apart. After all, if a fine-grained view about moral individuation is true, then the basic units of moral analysis might be very small and intimately connected. In that case, the basic units of moral analysis might not be the relevant ones for many agents in many contexts.

Consider an analogy. Even if, say, waves are the *basic* units of physical analysis, objects like tables and chairs are still the *default* units of physical analysis for many agents in many contexts. Granted, we might sometimes need to think about waves, for instance when doing physics. But we mostly need to think about these larger objects. Similarly, even if parts of individuals are the *basic* units of moral analysis, individuals can still be the *default* units of moral analysis for many agents in many contexts. Granted, we might sometimes need to think about parts, for instance when allocating responsibility in

cases involving multiple personalities. But we mostly need to think about individuals. As in the physics case, this is simply the best approach for us given our practical and epistemic circumstances.

But notice that on this distinction between basic and default units of moral analysis, what counts as a default unit of moral analysis is agent- and context-specific. And in a world of connected minds, the default units of moral analysis might shift. For instance, we can imagine scenarios where more individuals such as octopuses exist, such that parts are more salient. In these scenarios, moral analyses might shift in the directions of parts. Alternatively, we can imagine scenarios where more groups such as insect colonies exist, such that groups is more salient. In these scenarios, moral analyses might shift in the direction of groups. And of course, we can also imagine scenarios where both kinds of shifts occur at once, and in these scenarios, moral analyses might shift in both directions at once.

To make sense of these possibilities, consider a different analogy. For much of the twentieth century, international relations scholars treated nation-states as the primary units of political analysis. But in the second half of the century, increasing globalization led scholars to question that simplifying assumption. The result was a new global governance literature, characterized by the idea that any adequate analysis of international relations must make use of multiple units of analysis at once, including actors at the sub-national, national, and international levels. According to this approach, nation-states are still important units of analysis, but these other actors are important too. Indeed, in retrospect, actors at other levels were always important. Globalization simply made their importance more salient to us.

We believe that new mental connections will force the same kind of reckoning for interpersonal relations that globalization has forced for international relations. Ethicists currently treat persons as the primary units of moral analysis. But in the future, new mental connections will lead us to question that simplifying assumption. The result will be a new ethics literature, characterized by the idea that any adequate analysis of interpersonal relations must make use of multiple units of analysis at once, including actors at the sub-personal, personal, and interpersonal levels. Persons will still be important, but these other actors will be important too. And in retrospect, ethicists will realize, actors at other levels were always important. These mental connections simply made their importance more salient to us.

To be clear, we have not attempted to establish a fine-grained view in this section. Instead, we made a general distinction between fine-grained and course-grained views, noted a general advantage that fine-grained views have over course-grained views, and noted that these questions will become more important in a world of connected minds. We also made a general distinction between basic and default units of moral analysis, and we noted that even if individuals are commonly the default units of moral analysis at present, they might not as commonly be the default units *in the future*, and they might not be the *basic* units even now. Soon enough, moral analyses might need to consider parts and groups more than they currently do, in a way that clarifies the structure that morality has always had.

We should also reiterate that while our examples in this section focused on one kind of mental connection – the combined minds that occur when personalities join to form persons, and then persons join to form nations – other kinds of mental connection are possible as well. In a world of connected minds, not only can minds interact in a “nested” manner, with nations interacting as nations, persons interacting as persons, personalities interacting as personalities, and so on. But minds can also organize across “nests”. A person in one nation can directly interact with a person in another, a personality in one person can directly interact with a personality in another, and so on. We are only at the start of imagining how complex, layered, and dynamic our moral analyses will need to become to capture that reality.

5. Conclusion

Our aim in this paper has been to survey the ethical questions raised by connected minds, as well as to motivate a general approach to answering these questions. Plausibly, many minds are connected now, both in human and, especially, nonhuman cases. And many more minds might be connected in the future, as new technologies such as AIs and biological-artificial interfaces come online. We must therefore ask whether connected minds can have connected welfares, rights, virtues, and other such moral properties and relations. We attempted motivated, without fully establishing, the idea that they do. Connected minds

can be connected moral subjects, and when parts of individuals, individuals, and groups of individuals all have minds at the same time, they can all be moral subjects at the same time.

But as we have emphasized, this survey is meant to be the start of a conversation, or perhaps the continuation of conversations started in the collective agency and personal identity literatures, not the end of a conversation. Moving forward, we will need to ask many questions in a deeper manner. This includes descriptive questions. For instance, further cognitive science is needed to understand how connected minds function in cases involving conjoined twins, multiple personalities, distributed cognition, and so on. Further computer science is needed to understand how connected minds (will) function in cases involving AIs. And further philosophy of mind is needed to understand which theory of consciousness is true and, so, how expansive the scope of connected (conscious) minds might be.

This also includes normative questions. While some of the moral questions raised in this paper are familiar from the literatures on collective agency and personal identity, other questions are new or, at least, modified. What are the best views of moral individuation, both in the fine-grained category and in the course-grained category, and what follows from these views about the scope and scale of the moral population in a world of connected minds? What are basic and default units of moral analysis in such a world? How should we answer questions about fairness, justice, and equality in such a world? And how should we answer questions about creation ethics and population ethics in such a world? Is a relatively connected future population better or worse than a relatively unconnected one, and why?

No matter how we answer these questions, we should prepare for our descriptive and normative analyses to shift. There is no denying that in a globalized world, we need new frameworks to understand how we relate to each other within and across nation-states, and how we ought to do so. Similarly, there is no denying that in a world of connected minds, we need new frameworks to understand how we relate to each other within and across individuals, and how we ought to do so. And while we might disagree about, for instance, whether to double count shared welfare states or whether to attribute individual rights to moderately psychologically connected minds, we can all agree that these connections will change how we should relate to ourselves and each other in important and surprising ways.