Cambridge Quarterly of Healthcare Ethics (2021), 30: 4, 585–603
doi:10.1017/S0963180121000062

CAMBRIDGE UNIVERSITY PRESS

THE GREAT DEBATES

Should We Use Technology to Merge Minds?

John Danaher[1]* and Sven Nyholm[2]

[1]School of Law, National University of Ireland, Galway, Ireland
[2]Utrecht University, Utrecht, The Netherlands
*Corresponding Author. Email: johndanaher1984@gmail.com

John Danaher: Opening Argument in the Affirmative

In a 1981 article, the well-known philosopher of mind, Paul Churchland, presented a fantasy regarding the possible use of technology to fuse minds together. He noted that recent findings in brain science (well, recent in 1981!) suggested that the two hemispheres of the human brain were, in fact, independent cognitive systems that communicated via a high-fidelity channel known as the corpus callosum. Ordinarily, this resulted in a frictionless and advantageous synchrony between the two hemispheres. Under the right conditions, however, (e.g., experimental tests of those with a severed corpus callosum) the two systems could be pulled apart once again.

Churchland wondered: if the corpus callosum could accomplish frictionless integration inside one skull, could future technology accomplish the same thing between different brains located inside different skulls? And if it did, would this not be a wonderful thing? As he himself put it:[1]

> "*if two distinct hemispheres can learn to communicate on so impressive a scale, why shouldn't two distinct brains learn to do it also? This would require an artificial "commissure" of some kind, but let us suppose that we can fashion a workable transducer for implantation at some site in the brain…Once the channel is opened between two or more people, they can learn to exchange information and coordinate their behavior with the same intimacy and virtuosity displayed by your own cerebral hemispheres. Think what this might do for hockey teams, and ballet companies, and research teams!*"

Let us call this *Churchland's Fantasy*. Recent developments in brain-to-computer and brain-to-brain communication technologies suggest that this may not be a fantasy for too long.[2] What I want to consider in this debate with my colleague Sven Nyholm is whether we should embrace Churchland's Fantasy? Would it, in fact, be desirable to fuse minds together? I'm going to argue that it would be.

In making this case, I will be arguing against a long tradition of dystopian science fiction—think of the Borg in Star Trek—that suggests that the fusion of minds is undesirable. I will also be arguing against entrenched traditions of individualism and liberalism, which appear to be anathema to such fusion. So, my aims are somewhat modest. I will not be arguing that the fusion of minds represents the best possible future for humanity, or that all people should want to fuse their minds with one another. Instead, I will be arguing that the fusion of minds could be desirable under the right conditions and that we should not be too quick to rule it out.

I should add one interpretive note before I begin: all of the arguments I am about to discuss were developed in collaboration with Steve Petersen. Fusing our two minds together, we wrote them up in an article called "In Defence of the Hivemind Society," which was published in *Neuroethics* in 2020. So, Steve deserves at least part of the blame for what I am about to say.[3]

What are We Debating?

It will help if I clarify what is meant by the fusion of minds. The mind is the organ that controls how we interpret, experience, respond to, and plan our interactions with the world. It generates beliefs, desires, intentions, emotions, memories, and so forth. It is the seat of our consciousness—particularly our felt experience of the world—and the source of much of what is unique about us as persons. Ordinarily we assume that minds are associated with particular persons or individuals and, given mild physicalist assumptions, we associate minds with specific bodies and, more particularly, brains.

The fusion of minds occurs when two or more minds, minds that were previously associated with different persons and located in different brains, are brought together to form a single mind. Churchland's fantasy presents one example of how this might happen, but I assume that it is a more diverse and subtle phenomenon than his thought experiment might suggest. I do not believe, for example, that the fusion of minds is an all-or-none thing: I think it can come in degrees. Some mental processes can be fused while others are left separate. For example, two minds might share common memory systems or motivational systems, while retaining distinct decisionmaking systems. I also do not believe that the fusion of minds is an absolute or irreversible thing. In other words, in order for Ann and Bob to fuse their minds together, it is not necessary for them to bind together in an irreversible way to form an entirely new, third person, Carl. The fusion of minds can be more partial or occasional than that. They might join together for certain purposes or to experience certain things.

In short, I think the fusion of minds comes in a range of different forms and be classified along different dimensions. Two dimensions of mental fusion seem particularly important to the present debate:

1. *The rational dimension*: Two minds are fused together for the purposes of performing one or more cognitive operations associated with rational agency, for example, belief formation, intentional planning, and so forth.
2. *The phenomenological dimension*: Two minds are fused together for the purposes of sharing phenomenological experiences of and thoughts about the world.

These two dimensions form a space of possible forms of mental fusion.

Not all forms of mental fusion will be that philosophically interesting or challenging. Indeed, many incidental or occasional forms of rational fusion are probably commonplace: married couples, for example, are sometimes said to act like a single rational agent (for at least some purposes), as are some close siblings and friends. On the other hand, even mild forms of phenomenological fusion are not common at the moment and would seem to require some technological aid, of the sort envisaged by Churchland, before they could be said to arise.

The forms of mental fusion that are most philosophically interesting and challenging are those that involve high degrees of rational and phenomenological fusion—again, think of the Borg in Star Trek—where previously independent minds come together to form a single, enduring rational agent and share their experiences of the world. But we do not have to go all the way to that extreme to find interesting and challenging forms of mental fusion. Sharing experiences and thoughts over a period of days or months would be challenging to our sense of individuality too.

Is it technically feasible for us to achieve greater forms of rational or phenomenological integration? I am not sure. There are philosophers—Bryce Huebner springs to mind[4]—who claim that we are already able to achieve pretty significant forms of rational integration and there are tentative developments in brain-to-brain communication devices that might enable greater forms of phenomenological integration. I find these developments tantalizing and intriguing. But the technical feasibility of greater mental fusion is not my main concern. I am concerned primarily with the purely philosophical (axiological) question of whether it would be a good thing.

Five Arguments

I am going to present five arguments for thinking it might be a good thing. More precisely, I am going to present one argument for thinking that we should be open to the idea and four arguments for thinking that it would be desirable.

The argument from axiological openness

The first argument claims that we should be open to the possibility that the fusion of minds could be a good thing. Why? Well, we know for a fact that human moral systems and preferences change, to some extent, over time. There may be certain core values that are relatively universal and relatively fixed (e.g., the value of friendship, sociality, pleasure, and health)—although even this is controversial—but our attitudes toward those values, and the moral norms we associate with them, can and do change over time. Indeed, if we cast our gaze back but a few centuries, we can find our ancestors adopting and following moral norms that we now find abhorrent: wanton torture of criminals, slavery, denial of rights to women, and so forth. Where once we found those practices tolerable, perhaps even valuable, we now find them unpalatable and cruel.

Given this history of value change, it is plausible to suppose that our societal values will change again in the future and will change in a way that our offspring will look back on our time as one of great moral misalignment. At the moment, many of us, particularly those of us in Western, post-Enlightenment cultures, are deeply committed to values of liberal individualism: we think of the individual mind/person/body as the bearer of moral status and rights. We find any encroachment into the sacred space of individualism to be a threat to everything we hold dear. But this strong attachment to individualism is a relatively parochial and historically recent development. Ancient societies were far more communitarian (family or tribe-oriented) in their moral beliefs and practices, and many non-Western societies continue to this day to be more communitarian and anti-individualist (I will discuss the example of Buddhism in more detail in a moment).

The fusion of minds seems like a great assault on individualism. But given the history of value change, maybe my great-great grandchildren will look back on my attachment to individualism in the same way that I look back on my ancestors' attachment to torture and slavery? At the very least, we should be open to this possibility particularly if, as I shall now argue, the fusion of minds could help us to realize important values.

The argument from intimacy

The second argument claims that the fusion of minds could be a good thing because it could help us to achieve greater forms of intimacy. The starting presumption for this argument is that intimacy is a good thing. It is something that people desire and seek out in their lives. Indeed, for many people, their intimate relationships are often the most valuable and precious thing in their lives. Furthermore, philosophers often extoll the virtues of intimacy. Consider, for example, Aristotle's ideal of the virtue friendship, which requires a high degree of mutuality and intimacy between friends.

Intimacy can be achieved along a number of different dimensions (physical, psychological, etc.). As mentioned, it is already possible for two or more humans to bond together in a way that achieves a high degree of rational intimacy: they believe, plan, and act almost like a single agent. But our physical and phenomenological separateness prevents us from achieving yet higher forms of intimacy. We cannot feel what another person feels; we cannot see the world from their perspective. Using technology to achieve these additional forms of intimacy, perhaps in the manner envisaged by Churchland's fantasy, could be a good thing.

There is an obvious counterargument to this. Some people might argue that intimacy, as a good, presupposes the separateness of persons. In other words, what is good about intimacy is that two (or more) people become really close but remain separate. Indeed, you might even like to argue that it is conceptually confused to speak of intimacy arising from the fusion of minds: when two people are fused

together to form a single agent or phenomenological unit, they are no longer being intimate, they have instead transformed into a new, combined, entity.

There are two things to be said in response to this. First, at least as I envisage it, the fusion of minds need not always involve the elimination of separate individuals. This is already a feature of the forms of rational unity that intimate partners sometimes attain, and it may be possible to use technology to enable similarly temporary and partial forms of phenomenological unity: two lovers, perhaps, could share the same phenomenology for a couple of hours. Second, it is not obvious that intimacy requires separateness. There is nothing in either the concept or ideal of intimacy that seems to presuppose this. When people pursue intimacy, they are trying to break down the barriers between separate selves. The concept of intimacy itself seems to permit this activity being taken to its logical extreme (though, of course, this may have other side effects).

The argument from goal achievement

The third argument in favor of the fusion of minds claims that it could be a good thing because it could enable us to solve problems and achieve goals. Humanity faces a plethora of problems in the future: climate change, biowarfare, omnicidal terrorism, and other existential threats spring to mind. We need to solve these problems if we are to survive and thrive. Furthermore, solving problems and achieving goals, irrespective of its global benefits, is often thought to be a key element in the well-lived life. If I set myself the goal of solving some important intellectual puzzle, and I solve it, then this makes my life better than it might otherwise have been, all else being equal.

Individual humans have a limited capacity to solve problems and achieve goals on their own. Indeed, most of our successes and achievements as people come about with the direct and indirect support of others. As the anthropologist and psychologist Joseph Henrich has observed, the secret to humanity's success lies not in our individual intelligence but, rather, in our collective intelligence: our ability to collate information, share it, and manipulate it through shared language, books, cultural institutions, and others forms of collective endeavor. It is this that sets us apart. Furthermore, there is increasing evidence from science and engineering to suggest that most contemporary scientific and technological breakthroughs are not the work of lone geniuses but, rather, highly efficient and integrated teams of ever increasing sizes.

If problem solving and goal achievement are good things, and if collective teams are better at doing it than lone individuals, then this provides yet another argument in favor of the fusion of minds: why not use technology to ramp up our ability to collate, share, and manipulate information across separate minds? Why not use it to enable more efficient and effective forms of collaboration? Indeed, it may well be this argument that motivated Churchland's fantasy about the fusion of minds. Recall how he hinted at its benefits for "research teams."

Of course, one may object to this and argue that problem solving and goal achievement are not goods unless they are things that are achieved by individuals. *I* have to solve the problem and *I* have to achieve the goal in order for it to count. But this objection is suspect for two reasons. First, depending on the form of mental fusion being proposed, it may be possible for individuals to still share in the collective endeavor, much as they already do in team achievements (remember: mental fusion, at least as I envisage it, does not entail eliminating all separateness of individuals). Second, it is not clear that individualism is an essential part of the value of problem solving and goal achievement. There is often significant instrumental and intrinsic value to solving a problem, irrespective of who did it. For example, it was good that a COVID-19 vaccine was developed so rapidly. It matters less who developed it (and, of course, it is noteworthy that it was a large team, not an individual that did it).

The argument from moral imperfection

The fourth argument in favor of mental fusion argues that it could be a good thing as it would allow us to overcome certain obstacles to moral perfection. It is widely (though not universally) agreed among moral

philosophers that selflessness, altruism, and impartiality are key markers of moral thought and action. In other words, if we want to be more moral, we ought to be more selfless, altruistic, and impartial.

The problem is that individualism is an impediment to these forms of moral thought and action. This is a view that was famously suggested by Derek Parfit in his work on the reality of the self, but in suggesting this view, Parfit was merely echoing millenia of thought in the Buddhist tradition. It is well-known that Buddhist sects deny the metaphysical reality of the self. What might be less well-known is that some Buddhist sects, for example, the Abhidharma tradition, argue that commitment to the reality of the self-promotes moral defilements such as greed, jealousy, pride, and so forth. Consequently, they argue that it is both metaphysically and morally wise to rid ourselves of the illusion of the self.

The fusion of minds can help with this. By using technology to break down the barriers (illusory or otherwise) between individuals, we can address these moral defilements and come closer to the ideals of impartiality and altruism.

That said, I do not wish for this argument to be pushed too far. Individualism can be a source of moral defilements, for sure, but it can also be a source of moral virtue: courage, care, responsibility, and so forth. So, rather than saying that the fusion of minds is essential if we are to improve ourselves, morally speaking, I think it would be better to say that individualism is not the only moral game in town. Fusing minds and pursuing the idea of a hivemind society could also be compatible with moral flourishing, perhaps even a significantly improved form of moral flourishing.

The argument from meaning

The fifth argument in favor of the fusion of minds claims that it is both compatible with and possibly supportive of meaning in life. There are two elements to this argument.

First, there is the claim that it is compatible with the traditional accounts of what it takes to live a meaningful life. There are many traditional accounts of meaning. Some claim that achieving certain subjective states of mind is crucial (e.g., desire fulfillment and preference satisfaction); some claim that achieving certain objective ends is crucial (e.g., scientific insight and moral improvement); and some claim that a combination of both is crucial (e.g., being fulfilled by pursuing projects of objective value). It might be thought that these traditional accounts of meaning are incompatible with mental fusion: by eliminating the self we eliminate the person who is supposed to be living a meaningful life. But this does not seem quite right. For starters, as already noted, the fusion of minds does not necessarily eliminate separateness and, even if it did, one could argue that it results in the creation of a new subjective consciousness or agent that can live a meaningful life in the ordinary way. In addition to this, if we look at the main accounts of what it takes to live a meaningful life—subjective, objective, or a combination of both—it is not obvious that they require a specific individual to be having those experiences or achieving those ends.

Second, there is the possibility that mental fusion could enable us to achieve forms of meaning that would be inaccessible to us as individuals. I am thinking, in particular, here of accounts of meaning that claim that pursuing integration with something larger than ourselves is a source of meaning in life. Robert Nozick, for example, once argued that transcending our natural limits and pursuing integration with a larger organic unity was one pathway to meaning in life.[5] It is difficult to completely unpack this mystical ideal, but it seems to me that one plausible interpretation of it is that pursuing mental fusion—integration with other minds—is a pathway to meaning in life. If that is right, then we have another reason to favor mental integration.

Conclusion

This will suffice for my opening statement. There is more to be said, no doubt, and plenty of objections to be raised, but I hope this goes to show why I think using technology to fuse minds together could be desirable.

Sven Nyholm: Opening Argument in the Negative

Last year, John Danaher participated in an exchange that was similar to this one, where he discussed his 2019 book *Automation and Utopia*[6] with a mutual friend and collaborator of ours, Brian Earp.[7] In his introduction of Danaher, Earp said:

> One thing I like about your writing, John, is that you'll often explore extremely controversial topics, and you'll sometimes lead the reader into what can seem like extreme or troubled territory. But you will do it in a way that is calm and reasoned and often leads the reader by the end of the discussion [into] thinking "well, perhaps that idea isn't quite as absurd or shocking or bizarre as I initially thought it was."

As I have been preparing for this exchange, that statement about Danaher's work from Brian Earp has popped into my mind a few times. Except in this case, although Danaher's line of argument is clearly calm and reasoned, I am not sure whether I have ended up with the conclusion that Danaher's idea was not as absurd, shocking, or bizarre as I initially thought it was.

One reason for this is that I am still not sure that I fully grasp exactly what the idea of using technology to fuse minds involves, because it seems to me there are many different possible interpretations of what Danaher might mean. I hope that one of the things we can do in this exchange is to further clarify what exactly one might mean by this idea of fusing minds. This will be my first main topic.[8]

My second main topic will be Danaher's claim that we should be open to the idea that it might eventually come to seem desirable to us to fuse our minds with others, even if we do not yet think that this is a good idea, and even if this clashes with the individualism at the heart of a lot of our current thinking about what is good and important in life.

My third topic will be a brief—perhaps too brief—reflection on the arguments that Danaher puts forward in favor of thinking that it would be desirable to fuse our minds. I think my role in this exchange is to raise some skepticism about this—and I am indeed skeptical to some extent about this idea. But I also want to note right at the beginning that I am delighted to be discussing these issues with John Danaher. Even if I am not sure that I completely understand this particular idea of his, and even though I am also somewhat skeptical of it as I think it should be understood, I am an admirer of Danaher's work. I always enjoy engaging with it, even when—or perhaps especially when—I do not fully agree with the ideas he is arguing for.

Fusing Minds?

Danaher has a wonderful podcast—*Philosophical Disquisitions*[9]—that I am also a fan of. One of the things I like about that podcast is that Danaher often starts his interviews with his guests by having a discussion about how to understand the most basic ideas or concepts related to the topics covered in the different episodes.

Since our topic is whether we should fuse our minds, and it is not obviously clear what this might mean, it seems like a good idea to first reflect a little on what exactly it might mean. Is this a coherent idea to begin with? If not, it cannot be desirable to fuse our minds. And it cannot be that we should fuse our minds. At any rate, some ways of fusing our minds might be much less desirable than others. So, it is good to get clear on what we are talking about.

Danaher discusses two main ways of fusing our minds:

1. *Rational fusing*: Minds are fused for the purpose of performing cognitive operations associated with rational agency.
2. *Phenomenal fusing*: minds are fused for the purpose of sharing phenomenal experiences.

The first kind of mind fusing is less confusing than the second. When two or more people act together, they can form what is sometimes called a "group agent" or "corporate agent."[10]

As it happens, for example, Danaher and I—and Danaher, Brian Earp, and I—have written academic papers together, and thereby "fused our minds" in the rational way, for the purpose of writing those papers. Notably, those papers have had a somewhat different character than papers I have written together with others—such as the papers I have written with Lily Frank or the ones I have written with Steve Campbell. And they have also been different—and the writing process has been slightly different—from the work I have written on my own.

So, for me, the idea of rational fusing of minds makes sense, at least when one thinks of a domain-specific, reversible, and limited form of rational fusing of minds of the sort that is involved in group agency. In fact, I have even argued in some of my other work that we can engage in group agency together with robots that have a form of artificial intelligence minds (which are very different from human minds).[11] So again, for me, the idea of rational fusing of minds is not unclear. What is much less clear to me, however, and what I would like to ask Danaher to say a little more about, is the idea of phenomenal fusing of minds.

What I am particularly wondering about is how Danaher thinks about the relation between our phenomenal experiences and our bodies and nervous systems. It seems clear to me that people's phenomenal experiences are very dependent on their particular bodies and nervous systems—including what happens to their bodies and nervous systems in the particular environments they are currently in.[12]

So, if our minds are fused, then the experiential or phenomenal mind fusing would presumably have inputs from different bodies and nervous systems, which might react differently to the environments they are in. This could potentially lead to mixed and strange phenomenal states, if the idea makes sense to begin with.

For example, when I find a room that I am in to be too warm to be comfortable, my wife sometimes finds the temperature pleasant. If our minds were fused, would we experience the room temperature as both pleasant and unpleasant? Or would we experience it as somewhere in between pleasant and unpleasant? Or what if I am feeling hungry while my wife is not, will our fused minds feel both hungry and not hungry, or a state in between the two? There will be inputs into the fused mind from one body (mine) that point in one direction, and inputs from another body (my wife's) that point in another direction. And those bodies and nervous systems might respond differently to the same inputs (e.g., the room temperature). Or consider a case in which one person feels full or unwell, whereas others involved in the fusion do not feel full or unwell, but hungry or eager to do something. Again, one wonders what the resulting fused mind will feel or experience.

Could there be a way of editing out the unpleasant experiences, and have the fused mind only experience the pleasant feelings and sensations? If two people are on a boat together, and one is feeling seasick, could one switch to only experiencing the feelings of the other, who might not be feeling seasick? (When I have been on boats, I have sometimes wished that this were possible!)

These questions may all depend on a lack of imagination on my part. But what they are intended to bring out is that it is not so clear how the idea of phenomenal fusing of minds would work—especially not if the fused experience is going to have some unity to it.

The problem seems to be precisely that our experiences are very dependent on our bodies and brains and how they react to the environment and what is happening to us—and that people sometimes react in different ways to similar things. Or perhaps the people with fused minds are in different places, with very different inputs to their bodies—for example, one is in a very hot room, whereas the other is out in the cold. The resulting feelings and experiences would presumably be rather strange.

Should We Be Open to a Future in Which We Value Mind Fusion?

Danaher is assuming that because of our current individualism, many of us are unlikely to find the prospect of mind fusion desirable. I think Danaher is right about this. But he also suggests that because people's values change over time—for example, few people find things like slavery or torture acceptable these days, even though many people in the past did—we should be open to changing our values, so that we might come to find mind fusion desirable in the future. Is Danaher right about this? Should we be open to this possible change in our values?

This depends a little on what we mean by this openness. If it means that we should leave open the possibility that, as a matter of fact, most people might change their minds, then Danaher is surely right that we should not rule out this possibility. As he notes, there have been changes in people's values in the past. And so there will likely be changes in the future as well, some of which might seem quite radical from our current point of view.

In contrast, if we mean that we should think that it might be good or desirable to change our minds about what is desirable, then it is less clear to me whether Danaher is right. If some past values are undesirable from our present point of view (e.g., the values of those who valued a system of slavery), it seems that we should stick to our current set of values, and try to avoid reverting back to those earlier values. So, if certain potential future values people might come to have also seem undesirable from our current point of view, is it not also the case that we should stick to our current values, since from the point of view of these values, those possible future values are undesirable?

It seems to me that there are two main ways in which we can assess past values or possible future values: either from the point of view of particular values we currently have, or from the point of view of some bigger set of values we currently have.

We can ask, for example, if we should change our ways of valuing work—as Danaher does in his above-mentioned book *Automation and Utopia.* And when we ask this question, we can reflect on it from the point of view of particular other values we currently have, for example, the value we assign to friendship or to romantic relationships. Perhaps we could promote or honor those other values (friendship and romantic relationships) better if we valued work less.

Or we could ask whether changing the way(s) we value work might be supported by a whole host of our current values—not only from the point of view of how we value friendship or romantic relationships, but also from the point of view of how we value justice, virtue, knowledge, personal autonomy, and so forth.

In order for it to be true that we should welcome a possible change in whether we find mind fusion to be a desirable prospect or not, it seems to me that either this should be because there are certain current values from the point of view of which this would be very desirable, or—better yet—it should be that this would be desirable from a whole host of current values that we have.

The mere fact that our values *can change* and *have changed* in the past does not seem by itself to support the conclusion that it is a good idea to be open to changing our values again in the future. What it supports is rather a value-neutral prediction that it could happen that we change our values in the future with respect to what we think of mind fusion.

Is Mind Fusion Desirable?

The question above was about whether we should be open to changing our minds about whether to welcome mind fusion. That is a somewhat indirect question about mind fusion's desirability. But Danaher also offers a number of arguments directly in favor of the desirability of mind fusion. I will end my opening statement by briefly discussing those arguments. Or rather, I will make some general claims about two different philosophical perspectives on values that one can use when assessing Danaher's arguments.

The remarks I will make here assume that Danaher is right that we currently value individualism and the separateness of persons—and that we do so at the same time as we value the other things Danaher discusses, that is, intimacy, goal achievement, moral perfection, and meaningfulness in life. According to Danaher, fusing our minds can help to promote intimacy, goal achievement, moral perfection, and meaningfulness in life, although it might clash—at least to some extent—with the value of individualism.

It is interesting to think about these arguments in relation to two different ways of understanding the relations among values that are interestingly discussed by Ronald Dworkin in his book *Justice for Hedgehogs.*[13]

According to a first view, which Dworkin associates with Isaiah Berlin, values are wholly separate from each other. They can conflict with each other. We should interpret them in isolation from each

other. And, accordingly, there might be tradeoffs between different values. For example, according to Berlin's perspectives, there might be a tradeoff between *freedom* and *equality* in the political domain. And what we understand by freedom is independent from what we understand by equality.

According to a second perspective—the one that Dworkin himself endorses—values should form a "unity of value." This means that we should interpret individual values in the context of how they are related to our other values. Values should mutually support each other. And although there can be different ways of prioritizing values, there are never any true conflicts between the most important values in life.

This is a little bit like the so-called unity of the virtues thesis that was discussed by Aristotle and many ancient philosophers. Returning to the example of freedom and equality, Dworkin would say that in order to work out a good interpretation of what freedom is, we should, among other things, make use of our best interpretation of equality, and vice versa. In the end, freedom and equality are not conflicting values, but part of the overall set of values associated with justice.

Suppose now that we take the Dworkin-inspired view—as I am inclined to do. Suppose, also that we think that some form of individualism is one of the values we should have as our overall set of values. If we make these two suppositions, the arguments that Danaher makes in favor of mind fusion using other values might become problematic.

Those arguments might become problematic because our interpretations of values like intimacy, goal achievement, moral perfectionism, and meaningfulness in life should be made to harmonize with the value that we also place on individualism. And it can seem that if we do still value individualism, then fusing our minds with others—and giving up our individuality—may not be desirable, even if it could be seen as promoting some of those other ends. Or rather, it would not promote those other ends in the best possible ways, that is the ways compatible with the value of individualism.

Now, these are pretty abstract remarks about Danaher's arguments—I am not zooming in on particular arguments and assessing them in isolation in the remarks above. Rather, I am trying to diagnose what view about the relation among values Danaher might need to take in order to be able to make the arguments he is making, if he is among those who value individualism, as I am assuming that he is.

That is to say, it seems to me that Danaher needs to adopt an Isaiah Berlin-inspired view of the relations among values, rather than the Dworkin-inspired view that I myself prefer. Only then can he argue in favor of individualism-unfriendly conclusions while at the same time valuing individualism.

If he thinks that individualism is valuable—and he also thinks that our individualism should influence how we interpret other values—then arguing for individualism-unfriendly conclusions using other values is less open to Danaher as an argumentative strategy. That is, if Danaher wants to have a "unity of value," and he still values individualism, he might have a harder time arguing for the desirability of mind fusion than if he is less concerned with unity among his values.

I will not discuss all of Danaher's arguments, but just focus on one of them, to illustrate this point with an example. I will focus on the last argument about meaning in life.

I think Danaher is right that it is possible to associate living a meaningful life with being part of something bigger than oneself. But I also think this is but one and not the only aspect commonly and rightly associated with living a meaningful life. I also think that it is a plausible idea that living a meaningful life involves living a self-directed life in which one enjoys personal autonomy. In fact, another friend and colleague of mine—Jesper Ahlin Marceta[14]—goes so far as to argue that leading a self-directed life is the very essence of leading a meaningful life.

I think that Ahlin Marceta is exaggerated in his very strong individualism about meaning in life. But I think that he is right that one key part of living a meaningful life is living a self-directed life. So, when we understand what it is to be part of something bigger than us, I think we need to interpret this idea in a way that makes it compatible with at the same time leading a self-directed life.

I doubt that we could do this in a sufficiently clear sense if we have fused our minds with the minds of others. But then again, this depends on what we mean by fusing our minds with other minds. And like I said above, I am still not sure I fully understand what Danaher has in mind. So, I think that this is a good point at which to hand things back to Danaher.

But to summarize: first, I find the idea of rational fusion of minds much clearer than the idea of a phenomenal fusion of minds. Second, I think that although we can predict that human values will change in the future, it is less clear that we should welcome those changes unless we can give arguments for this based on our current values. And finally, I think that if we commit to what Dworkin calls the ideal of a unity of value, and we have individualism as one of our key values, then some of the arguments Danaher offers in favor of mind fusion might lose some of their force.

John Danaher's Rebuttal: In Defence of Mental Fusion

I would like to thank Sven for his response to my opening statement. Sven kindly broke down his response into three main criticisms of my opening arguments. I will respond to each of these criticisms now.

Criticism 1—The Concept of Mental Fusion is Unclear

Sven's first criticism is that the coherence or meaning of the proposition we were debating is unclear. What does it mean to merge or fuse minds together? Sven thought the idea of rational fusion made sense insofar as it seems we do this in a temporary and ad hoc form as is. I am glad that he agrees with me on this point and I do think this is right: we form temporary but often meaningful rational bonds with other people on an ongoing basis. Perhaps this is not that interesting, but I would add that technology, particularly cybernetic prosthetics and brain-to-brain communication could make a more sustained and longer lasting form of rational fusion possible. I think this would be an interesting and significant thing in itself and I think there is a case for using technology to make it more possible that follows the structure of my opening argument.

However, setting that aside, Sven's main point was that the concept of phenomenal fusion is unclear and possibly even absurd or nonsensical. One reason why it seems like a strange notion is because our experiences are so intimately linked to our position within certain bodies in space and time. I experience the sensation of the keyboard at the tips of my fingers, but not at the tips of your fingers. I experience my own sense of hunger but not your sense of hunger (or satiation, as the case may be). To merge these distinct phenomenologies, intimately linked, as they are, to the one body, may be impossible or incoherent. At the very least, it could generate an intuitively strange or radically different set of experiences of the world.

There are several things I want to say in response to this criticism. An obvious point is that just because it seems intuitively strange to imagine the phenomenology that might emerge from linked minds does not mean that it is impossible or absurd or even incoherent to link minds. There are many things that are possible in the real world that seem intuitively strange. Physical laws such as quantum entanglement are, for example, contradictory to folk physics and folk understanding. That said, if something seems intuitively odd to us -- as Sven seems to think phenomenal fusion is—it may affect our capacity to evaluate it from a moral perspective. This might, to some extent, undermine my argument. In this sense, intuitive strangeness could be relevant to our debate. If we cannot meaningfully assess what it might be like to be a merged mind, then perhaps we cannot argue about its axiological desirability. But, of course, here we encounter one of the perennial puzzles in the philosophy of the mind: phenomenal experience is inherently private. We can never know "what it is like" to be another person? Does this make it irrational to make educated guesses?

In any event, I am not sure that it is intuitively strange or leads to the kind of oddness or incoherence that Sven imagines. It seems to me that Sven might be a victim of the myopia that comes from having a mind that is located in a single body—call this "the single body fascism."

I think we can overcome this myopia. It is not that unusual or strange to imagine phenomenal fusion; after all our brains manage to integrate sensory signals from multiple sources, different limbs, different sensory modalities into a single phenomenal experience of the world. Our brains do this every day. In

addition to this, sometimes over our lifetime we lose some of these sensory signals and inputs. This can affect our phenomenal inputs and will require some kind of phenomenal readjustment. For example, amputation, blindness, and being in a sensory deprivation tank are all things that affect our phenomenal inputs and that in turn affects our phenomenal experience of the world around us. Nevertheless, we adjust to varied phenomenal inputs, sometimes, admittedly, not that well, but given that we can do this despite being located in just one body, it is not immediately or intuitively obvious that we could not readjust to the addition of new sensory inputs from different bodies. Indeed, there are preliminary technological innovations that create new sensory inputs to which our brains must adapt: neuroprosthetics and virtual reality headsets are just two examples.

I think that some of the weirdness or strangeness that Sven points out comes from thinking that this would result in some superposition of incompatible mental states. This is the concern that is at play in his examples of experiencing hot and cold at the same time or being full and being hungry. But I think that is not likely to happen in practice. It is more likely that the integration of different phenomenal inputs would lead to a single new phenomenal state which could either be a blending of the two original states or a new emergent phenomenal state. Consider, for example, what it might be like to have one single experience of having four legs and not two separate experiences of having two legs (incidentally this is not unprecedented: some cases of conjoined twins have involved the sharing of bodily organs and systems).

Another point I would make here is that phenomenal or sensory switching between different bodies might be possible already and I think this can count as a kind of phenomenal fusion. For example, it is possible to get sensory input from a robot arm or prosthetic to feed into one person's brain so that they can experience that prosthetic as part of themselves. It is also possible to wear virtual reality (VR) headsets to give someone the sense that they are occupying someone else's body. Thomas Metzinger and Olaf Blanke have performed experiments on this. These experiments suggests that our "phenomenal self-models" can be manipulated with VR in a way that radically shifts our phenomenal self-perception.[15] My contention is that if there is a lot of switching back and forth between two or more peoples' sets of sensory inputs, using VR or augmented reality (AR) headsets for example, they could share an ongoing set of experiences of the world in phenomenally rich way that is quite different from the way we ordinarily experience the world. I would view this is a form of phenomenal fusion.

I will concede, however, that some forms of phenomenal fusion might turn out to be physically or biologically impossible. For example, it may well be the case that our brains cannot cope with lots of different sensory inputs; it could lead to sensory overload. But we cannot really know the limits until we try. If it does turn out that some kinds of phenomenal fusion are impossible, we can just focus on the kinds that are possible and run the argument from there.

Criticism 2—*We Cannot be Too Axiologically Open-Minded*

The second aspect of Sven's critique was in relation to the argument from axiological openness. This argument held that we should be open minded about future value change, including a shift away from individualism toward something like the shared-mind ideal. Sven argues that this is problematic because changes in value are only desirable from the perspective of current values.

In reality, Sven and I probably do not disagree on this particular matter. The four arguments I defended after the axiological openness argument are intended to provide reasons for being open minded about mental fusion that are based on current values, namely: intimacy, problem-solving ability, achievement, impartiality, and meaning of life. The purpose of the openness argument is simply to address a bias that people might have toward a certain cluster of values that seem to support strong individualism. It seems to me that favoring this cluster of values is historically contingent and we may come to regret it in the future. Humans find many things valuable and we can attach to different clusters of values over time. I think it is possible that we could attach to a cluster of values that supports mental fusion.

The axiological openness argument is intended as a kind of mild existential shock therapy to highlight the contingency of our current way of life and to suggest that we should be willing to open ourselves to a panoply of other choices as to what constitutes a valuable way of life.

Criticism 3—Values Must be Unified not Radically Plural

This brings me to the final set of arguments that Sven offered, relating to Ronald Dworkin's claims about the unity of value. If I can paraphrase that myself, I take it that Sven is arguing that all values should form a unity of value; in other words, although values may appear to be plural, they must ultimately be reconcilable with one another.

This creates a problem for my argument because if individualism is one of our current values, then it has to be reconcilable with all the other values. There can be no tradeoff between individualism and these other values. In other words, it has to be reconcilable with intimacy, goal achievement, and so forth. The problem with my case for mental fusion is that it assumes that individualism can be, or should be, traded off against these other values. So, my argument is not compatible with this unity of value thesis.

I have a number of responses that I can make to this argument. One possibility, which Sven himself hints at, is that I could just embrace an alternative view of value pluralism that disagrees with the Dworkinian view. For example, I could embrace something like Isaiah Berlin's view of value pluralism. I could hold that there is strong pluralism to our values; that at least some of the values that we have are incommensurate and possibly incompatible. Hence, there may be strong conflict and tradeoffs to be made between values. It is then up to us to pick between different constellations of values. If favoring one set of values—such as intimacy, goal achievement, problem solving, altruism, and so forth—requires us to sacrifice individualism then that is an acceptable choice that we might make.

Another response that is open to me is that I could just reject individualism as a value or argue that it is not an important value or foundational value. Perhaps, for example, it is a convenient label for a cluster of more foundational values. It is also possible that I could embrace a reductive form of pluralism that maintains that there is some kind of single underlying dimension to value, one that ultimately sustains a set of plural, but not incompatible, values but that is, in turn, not compatible with strong individualism. This would, in a sense, be a form of unity of value, just one that does not require fealty to individualism. For example, Robert Nozick's organic unity thesis, which I mentioned in my opening speech, supports a kind of weak pluralism about values that holds that all values are ultimately reducible to the pursuit of greater organic unity. I have already suggested that this supports my view.

Finally, I can also just reject the notion that mental fusion necessarily involves a tradeoff between individualism and other values. If the fusion between minds need only be partial or temporary, it could be the case that it just involves a reprioritization of values, a favoring of pure intimacy over individuality rather than an abandoning of the latter, and that might be compatible with the Dworkinian view.

I am not sure about the deeper philosophical issues here. I am agnostic, at the moment, between unity of value and strong pluralism, but it seems that the case for greater mental fusion is robust enough to work with either thesis.

Sven Nyholm's Rebuttal: Is Mind Fusion Desirable All-Things-Considered? Lingering Questions about Agency, Potential Risks, and Personal Identity

I am grateful for the opportunity to engage in this debate with John Danaher on whether we should use technologies to fuse minds. And I would like to start this rebuttal by thanking John for engaging so carefully and thoughtfully with the points I made in my opening statement. Whether or not we will ultimately end up "being of one mind" regarding the desirability of fusing minds, this is certainly a fascinating discussion. We are in perfect agreement on that point. But I remain puzzled about how best to understand the idea of mind fusion. Below, I raise some lingering questions about agency, potential risks, and personal identity.

How Many Agents?

One thing I think Danaher is still being a little unclear about is whether mind fusion would result in one agent with one big hive mind, so to speak, or whether there would be two or more agents, who somehow share or participate in a communal hive mind. Perhaps Danaher has not come down firmly in favor of either option because he thinks that both are possible. Perhaps he thinks that, depending on what kind of mind fusion is at issue, there are different possibilities. Notably, Danaher discusses the possibility of temporarily fusing minds and then unfusing them again. Perhaps the best way of interpreting that idea is to say that mind fusion at least sometimes allows for the possibility of retaining one's own individual identity and of being one's own agent, even though one sometimes temporarily becomes part of a fused mind.

When it comes to the earlier-discussed idea of group agency, however, the suggestion is usually that different agents can come together to form a new, corporate agent (e.g., a business corporation, a sports team, a university, or a church). This corporate agent might have aims of its own, sometimes performing actions that cannot be attributed to any of the particular members making up the group agent.[16] So, if we are to use the idea of group agency as a model for what Danaher calls rational fusion, this would suggest that rationally fused agents form a new agent. Similarly, if we think of phenomenal fusion—that is the fusion of people's experiences and sensations—this would also most easily be interpreted as the formation of a single agent with a single mind, who would have a body that is in effect scattered in space and time (i.e., the different bodies whose minds are fused into this hive mind).

Yet, as I noted above, Danaher appears to be open to the possibility that there is not one, single agent that is the result of mind fusion. Rather, there are different agents—who are separate from each other—with a shared mind they all have access to. I bring up these two different interpretations—the interpretation that one mind and one agent is created, and the interpretation that one mind but many agents are created—because I think that each might have their own problems. This can be seen as creating a dilemma of sorts when we start thinking about possible risks involved in the prospect of fusing minds. That is, there could be distinctive problems related to each of these two possibilities regarding what the result would be in terms of how many agents there are if different minds are fused. I will discuss this now, as I turn to the topic of whether mind fusion is merely desirable in certain ways, as Danaher sees things, or whether it is desirable all things considered.

Mind Fusion and Potential Risks

Since Danaher's aim is to argue that mind fusion is desirable, it is not surprising that he focuses primarily on potential benefits, rather than on possible risks or harms. It might very well be that mind fusion is desirable in certain respects. Danaher has made a strong case in favor of the idea that it is desirable in the various respects that he discusses in his opening statement. However, Danaher also discusses the claim that we *should* fuse our minds. This proposition—that we should fuse our minds—suggests something stronger than that there merely are certain respects in which mind fusion might be desirable. It suggests that, all things considered, it is on the whole more desirable than undesirable to fuse minds.

Importantly, in order to come to a considered judgment about whether mind fusion is desirable all-things-considered, we need to reflect, not only on potential benefits, but also on potential risks. So, here I want to bring up some risks it would be good to think more about before we conclude that mind fusion is desirable (or undesirable) all things considered. These risks depend a little bit on whether we think there is one agent with one mind when minds are merged, or whether we think there are multiple agents, who somehow have access to a shared hive mind.

Let us start with a risk we can imagine if the hive mind is associated with one hive agent, so that the individual humans who join the hive mind lose their own identity as separate agents. If this happens, a worry might be that members whose minds are thought to not make any valuable contribution to the overall hive mind might be excluded. Perhaps some people are thought to have minds that just do not add

anything of value. Or perhaps they are thought to add discomfort of some sort to the overall agent with the fused hive mind. Danaher mentioned—in jest—the idea of "single-body fascism" above. Well, there might also be a worry about multibody fascism if the resulting hive agent starts discarding or "cleansing" members perceived as undesirable.

If we instead envision the fused hive mind as not forming a single agent, but as rather being a shared mind that different agents have access to, then other worries might arise. One thing that one cannot but worry about is what happens to the privacy of the different members of the hive mind. If others, who are separate from us, have some form of direct access to our minds via the hive mind technology, this can sound like a privacy nightmare. In the ethics of technology, worries about threats to privacy are very common.[17] Here, with the idea of technologies used to create a shared mind, it seems that these privacy worries would be turned up to a whole new level. This would especially be so if the technologies used for creating the hive mind(s) would be owned by private corporations looking to profit from the creation of these hive minds.

Speaking of the plural of "hive mind" and speaking of worries from the ethics of technology, another risk that should be considered—and which is also very familiar from the ethics of technology—is the worry about polarization and filter bubbles/echo chambers.[18] Supposing that not all human beings join together in one huge fused mind but rather that there would be two or more fused hive minds, a key worry would be that these resulting group minds would be polarized from each other, potentially becoming very set in their ways and mentally shielded off from the other hive minds. Perhaps with hive minds there would be a very extreme form of "group think" resulting, and this might lead to polarization on steroids. This seems like another possible risk to take into account when we think about whether fusing our minds with others would be desirable all things considered. This polarization risk would seem to be there and worth worrying about independently whether we think of hive minds as involving only one agent or whether we think of hive minds as involving more agents, who somehow share a mind.

Puzzles about Personal Identity

The above-mentioned potential risks related to mind fusion are examples of things we should take into consideration when thinking about whether mind fusion created by advanced technologies would be desirable all things considered. Like I said above, since Danaher was mostly focusing on possible benefits of mind fusion in his opening statement, I wanted to bring up some possible risks and potential problems here, to highlight that we should also take the possible negative outcomes into consideration, and not only the possible positive aspects of mind fusion.

I think it is important to consider the kinds of worries I have brought up in the section above. Moreover, there might be other worries that I have overlooked in my brief discussion earlier. In general, I think it would be good to think more about possible risks related to mind fusion. But at the same time, I want to end this rebuttal on a less negative note. So, the last thing I will do here is to bring up another question that came to mind for me as I was reflecting on Danaher's intriguing ideas about mind fusion. I was prompted to think of the issue that I will now bring up because Danaher mentioned Derek Parfit in his opening statement and because I was trying to think about different interpretations of what the idea of fusing minds might involve.

One interesting question that arises—and that is closely related to Parfit's well-known theory about what personal identity over time consists in[19]—concerns the issue of whether fused minds would come to form memories based on inputs from the experiences of the different people whose minds are fused. Might it be that, over time, a chain of memories and other forms of psychological connectedness and continuity might be created in the fused minds? Perhaps new generations of people could gradually come to merge their minds with the fused hive mind, adding in new memories and helping to bring old memories into the future, and more generally helping to create psychological connectedness and continuity within the hive mind over time. Perhaps old members of the hive mind might die or for other reasons stop being part of the hive mind. And yet memories based on their experiences, or other forms of psychological continuity from them, might continue on into the future in the hive mind.

According to Parfit's above-mentioned theory of personal identity over time, personal identity depends on chains of psychological continuity and other forms of psychological connectedness.[20] If this is right—and a hive mind might come to have memories of experiences had by past members, and create psychological continuity into the future with new members—what would this mean for the identity over time of the hive mind? Perhaps we would have a striking situation where the hive mind has an identity of its own, which is able to survive the members (the people) whose experiences, thoughts, and so forth are fused in the hive mind.

One is here reminded of the saying "the king is dead, long live the king!"—that is the idea that a type of immortality could be associated with a role, which different human beings could step into and play over time. The same idea is sometimes discussed in relation to corporate agents as well. A firm, a sports team, a church—or whatever it might be—can continue existing over time even though all the members are replaced with new people. So long as there is the right kind of continuity in the organization, the corporate entity can exist indefinitely.

It is really fascinating to think in these terms when we reflect on what the idea of mind fusion might amount to. What should we say about the identity over time of a hive mind of the sort Danaher envisions? Could it potentially enjoy a form of "immortality," whereby a new agent or personal identity is created that could in theory persist over time indefinitely? This might depend on all sorts of things, including whether we should accept a Parfit-inspired theory of what personal identity over time amounts to.

I wanted to end with this fascinating issue, rather than the ethical worries about the technology-generated hive mind briefly discussed above in order to make it clear that although I think that Danaher's suggestion does raise important ethical concerns, it also brings up some deep and truly intriguing philosophical questions. It raises questions about how to think about agency and personal identity in future scenarios in which people might relate to each other in radically new ways that we cannot fully grasp or easily comprehend from our current point of view.[21] So, to conclude, I do remain skeptical about the overall desirability of mind fusion. And I am still unsure about how best to understand this idea to begin with. But I fully agree with Danaher that this is an absolutely fascinating topic to reflect on.

John Danaher: Sven Nyholm's Best Argument

It is always a pleasure to debate questions like this with Sven. He is the perfect example of a "good faith" interlocutor, someone who takes the alternative point of view seriously and highlights weaknesses within it, without needless point-scoring or ad hominems. I am sure that he must occasionally find my views absurd, but I never get the impression that he finds me absurd and that is always reassuring. Furthermore, in presenting his skeptical doubts about my arguments he helps to rein-in my more outlandish speculations.

I think Sven's best objection to my position is his first one, that there is something slippery or uncertain about the proposition that we are debating. Although I believe that there is broad agreement in philosophy about what a mind is (i.e., the concept is not essentially contested as many others are), there is no doubt that minds have many parts. There is, for example, the part that is responsible for forming beliefs, the part that generates desires and the part that plans actions. There are also the parts of the mind that are unconscious and the parts that are conscious.

Each of these parts is likely to have many subcomponents in its own right. For instance, different subparts might be responsible for producing different kinds of desire, or forming different memories, or experiencing different aspects of the world. To the extent that the mind simply is the brain, we know that this must be true. Functional neuroanatomy suggests that there are specialized subprocessors within the brain that are responsible for different mental functions. These may communicate rapidly and fluidly with one another, such that there is some degree of "global" integration, but they are also dissociable. Indeed, experiments highlighting the dissociation of different mental functions were what prompted *Churchland's Fantasy*, which was the jumping off point for this debate.

So, when we debate whether it is desirable to fuse minds together, what exactly are we debating? The fusion of all of these parts and subparts or just some of them? What would the resulting concatenation of

mental functions and events be like? Can we really speculate about this in a meaningful way? Sven is right to urge greater clarity on these matters. When I reflect on my own position in this debate, I have to accept that I do sometimes slide back and forth between more and less extreme versions of mental fusion. Sometimes, I am talking about the fusion of one or two mental functions (e.g., intentional planning in the case of rational fusion or temporary experiential sharing in the case of phenomenological fusion), but other times, I am probably appealing to a more wholescale form of mental fusion (e.g., in my comments about meaning in life and pursuing organic unity).

In this respect, and building on Sven's critique, I think it is possible to accuse me of committing the "motte and bailey" fallacy. First described by Nicholas Shackel, this is an informal fallacy that is common in debates about controversial topics.[22] It is named after a medieval system of fortification in which a village would be constructed on a low-lying piece of land (the bailey) and a fortified tower on a higher piece of land (the motte). If the village was ever under attack, people could retreat from the poorly defended bailey to the better-defended motte. Shackel argued that philosophers occasionally do the same thing when defending controversial but vague propositions. They start out with what seems like a bold and outrageous claim and when this is attacked they retreat to a more specific and less outrageous claim that is similar in nature. Perhaps I am doing something similar by proposing a bold and outrageous claim (we should fuse minds) but then retreating to a more specific and less outrageous variation on this (we should temporarily share experiences) when challenged?

In my defence, when Steve Petersen and I originally wrote about the notion of a "hivemind" society, we viewed the vagueness of the proposition as a feature and not a bug in our argument. Our goal was to argue that there is a sliding scale of different forms of mental fusion, some of these are more immediately desirable than others but with each step we take along this sliding scale the more likely it is that more extreme versions of mental fusion seem desirable. That said, in order to take this debate forward, and make a more robust case for mental fusion, it might be worthwhile being a bit more specific and focus on specific forms of mental fusion.

I will also add, as a final note, that Sven's comments about the risks inherent in mental fusion are well-taken. Of course, when considering a possibility like this we should factor in the benefits and the risks.

Sven Nyholm: John Danaher's Best Argument

John Danaher has offered many good arguments and put forward many highly interesting and provocative claims in the course of this debate. As I said in the introduction when I was quoting Brian Earp, Danaher tends to explore what can seem like "extreme and controversial topics," and I think this exchange has helped to illustrate that point. After all, the claim that we should consider giving up our individuality and merge our minds with others is likely to strike many as an extreme and controversial idea. But as I also said, at the end of an intervention by Danaher—whether it is about the future of work,[23] the moral status of robots,[24] or the threat of rule by algorithms[25]—one definitely comes away with useful insights and with a sense of having learned something new and important. This exchange is no exception to that rule.

The very last thing I will do in this exchange is to comment on what has been the best argument—or, rather, the best arguments—from Danaher's side in this debate. Before I do so, however, I would also like to point out that simply raising this provocative topic of whether we should use advanced technologies to merge our minds, presenting it in such an interesting way, and also offering such a compelling case for what initially seems like a slightly crazy idea is already an achievement on Danaher's part. But let me now comment on what I think were Danaher's strongest points along the way.

As I see things, the perhaps most compelling reason Danaher has offered in favor of mind fusion is encapsulated in his argument that this might help to facilitate problem solving. I agree that when people do things together, they can usually achieve much more than when they try to do things completely on their own. Indeed, it is hard to achieve much at all on one's own.

What are presented as the accomplishments of lone geniuses are really, in all honesty, usually the outcomes of processes involving many different people, including those who inspired the supposed lone

genius or provided the source materials he or she was drawing on. This point is nicely made, for example, in Emma Smith's wonderful *This is Shakespeare.*[26] Throughout her discussion in that enjoyable book, Smith investigates on one hand Shakespeare's sources and help he received from his contemporaries; and on the other hand, explains how readers, performers, directors, and other interpreters help to bring Shakespeare's works to life. If not even Shakespeare can create masterpieces all by himself in isolation, but works best with the help of others, this surely applies to the rest of us as well.

Creating striking works of art could perhaps be one thing that we could do if we fused our minds in the way(s) that Danaher imagines. But more importantly, it might help us to solve some of the most pressing issues and biggest problems that we are facing as a society today and in the near and distant future. Getting a handle on what to do about human-created climate change, for example, and coming up to a solution to that problem—including the issue of how to make people respond more decisively to this enormous threat—is something we might need to fuse our minds to create a super-mind in order to be able to do. In my mind, the idea that mind fusion could help us to tackle important problems better and achieve greater things is the strongest argument Danaher has presented in favor of mind fusion.

At the same time, I also see some clear appeal in Danaher's claim that joining together with others through mind fusion could add to the meaningfulness of our lives. Perhaps I am led to reacting to that argument of Danaher's by looking around and seeing so much division in the world right now, which can seem like the opposite of meaningful ways of relating to one another. Not to become overly political, but the highly divisive tribalism in current politics in the United States, for example, strikes me as not only being meaningless, but also as being the polar opposite of meaningful: what you might call "anti-meaningful."[27] It would be more meaningful if people could join together; and perhaps mind fusion would be an even more meaningful way of doing this. At the same time, although, I do still worry about the threats that mind fusion poses to the value that most of us place on individualism. And I think that we typically do—and rightly do so—understand the ideal of living a meaningful life as making room for the values we associate with individualism. However, as I have said here, I do see Danaher's claims about how mind fusion could promote at least one aspect of meaningfulness in life as one interesting, if not strong, argument in favor of that idea. In the end, I remain conflicted about this suggestion of Danaher's.

The last thing I will mention is that I also find Danaher's insistence that we should take up an attitude of axiological openness to be a strong part of the overall perspective that he has presented. Again, when one looks around and sees how unwilling to consider other people's perspectives some people seem to be—and how prone people are to coming to accept conspiracy theories that demonize others and their perspectives—Danaher's call for axiological openness looks very good in comparison. Personally, I do not think we should be too open to changes in values, as I argued in my opening statement. I think we need to assess potential future changes in our values from the point of view of either individual values we currently hold or sets of values we now endorse. But this, perhaps slightly more conservative version of axiological openness—what we might call "cautious axiological openness"—is perhaps not very far from what Danaher has been meaning to defend all along. Indeed, in his rebuttal in response to my opening statement, Danaher seems to be in partial, if not full, agreement with me about how we should temper our axiological openness by relating it to values that we endorse here and now and that can thereby guide our assessment of future possibilities.

It seems to me that doing what Danaher is doing in his work—namely, presenting controversial ideas and arguing for them in what Brian Earp has characterized as a "calm and reasoned" way—is a good way of promoting axiological openness. So, I would like to end by noting that although I think that the argument relating to problem solving is the best argument Danaher has presented in favor of mind fusion, his overall approach provides a strong case in favor of axiological openness. That is to say, in this and other interventions of his, Danaher has given us and is continuing to give us powerful reasons for both considering and taking seriously some striking ways, in which we might want to change our values as we move into the future.[28]

Notes

1. Churchland P. Eliminative materialism and the propositional attitudes. *Journal of Philosophy* 1981;**78**(2):67–90, at p 88.
2. For a popular introduction see Nicolelis M. *Beyond Boundaries: The New Neuroscience of Connecting Brains with Machines and How it Will Change Our Lives*. New York: St Martin's Press; 2011.
3. See Danaher J, Petersen S. In defence of the hivemind society, *Neuroethics* 2020. https://doi.org/10.1007/s12152-020-09451-7
4. Huebner B. *Macrocognition: A Theory of Distributed Minds and Collective Intentionality*. Oxford: OUP; 2014.
5. Nozick R. *Philosophical Explanations*. Cambridge, MA: Harvard University Press; 1981, at chaps. 5 and 6.
6. Danaher J. *Automation and Utopia*. Cambridge, MA: Harvard University Press; 2019.
7. A recording of that event is available here: *Automation and Utopia: John Danaher in Conversation with Brian D. Earp*; available at https://www.youtube.com/watch?v=lTqo_zVMRgI&t=268s (last accessed 4 Feb 2021).
8. One worry about John's idea of using technology to merge minds might be that it is unclear what kinds of technologies could be used to accomplish this feat. Here, however, I will mostly set technological worries aside and instead primarily focus on more conceptual and evaluative/normative issues.
9. https://philosophicaldisquisitions.blogspot.com/p/podcast.html (last accessed 4 Feb 2021).
10. See, for example, List C, Pettit P. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press; 2011.
11. Nyholm S. *Humans and Robots: Ethics, Agency, and Anthropomorphism*. London: Rowman & Littlefield International; 2020. See, in particular, chaps. 3 and 6.
12. See note 11, Nyholm 2020, at chap. 6, p. 146
13. Dworkin R. *Justice for Hedgehogs*. Cambridge, MA: Harvard University Press; 2011. Part of my inspiration for bringing up Dworkin's discussion in this context comes from reading Anna Melnyk's paper-in-progress "An Interpretation of Value Change," which also relates Dworkin's ideas to the issue of value change influenced by technological developments, particularly potential value changes related to worries about climate change and new energy sources.
14. Ahlin Marceta J. An individualist theory of meaning. *Journal of Value Inquiry* (in press).
15. Blanke O, Metzinger T. Full-body illusions and minimal phenomenal selfhood. *Trends in Cognitive Sciences* 2009;**13**(1):7–13. https://doi.org/10.1016/j.tics.2008.10.003; these experiments and others are discussed, along with their philosophical implications, in more detail in Metzinger T. *The Ego Tunnel: The Science of the Mind and the Myth of the Self*. New York: Basic Books; 2009.
16. See, for instance, List C, Pettit P. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press; 2011.
17. See, for example, Veliz C. *Privacy is Power*. London: Penguin; 2020.
18. See, for example, Pariser E. *The Filter Bubble: What the Internet is Hiding from You*. London: Penguin; 2011; Lynch, MP. *The Internet of Us: Knowing More and Understanding Less in the Age of Big Data*. New York: Liveright; 2017.
19. See part 3 of Parfit D. *Reasons and Persons*. Oxford: Clarendon Press; 1984.
20. See note 19, Parfit 1984.
21. Perhaps we should think of mind fusion as a "transformative experience" in L.A. Paul's sense, which we cannot rationally evaluate before we have experienced it ourselves. See Paul LA. *Transformative Experience*. Oxford: Oxford University Press; 2014.
22. Originally described in Shackel N. The vacuity of postmodernist methodology. *Metaphilosophy* 2005;**36**(3):295–320; Shackel referred to it as a "doctrine" rather than a "fallacy."
23. See note 6, Danaher 2109.

24. Danaher J. Welcoming robots into the moral circle: A defence of moral behaviourism. *Science and Engineering Ethics* 2020;**26**(4):2023–49.
25. Danaher J. The threat of algocracy: Reality, resistance and accommodation. *Philosophy and Technology* 2016;**29**(3):245–68.
26. Smith E. *This is Shakespeare*. London: Penguin; 2019.
27. Together with Steve Campbell, I have elsewhere explored the worry that modern life involves various different problems that threaten to make our lives the opposite of meaningful or "anti-meaningful," as we put it, for example in Campbell SM, Nyholm S. Anti-meaning and why it matters. *Journal of the American Philosophical Association* 2015;**1**(4):694–711; Nyholm S, Campbell SM. Meaning and anti-meaning in life. In: Landau I, ed. *Oxford Handbook of Meaning in Life*. Oxford: Oxford University Press; in press.
28. Many thanks to John Danaher and Tomi Kushner for helping me during the process of putting together my contributions to this exchange. In creating the written version of this debate, I have also been helped greatly by the discussion session that we had at the Neuroethics Network event in February 2021, which was chaired by David Lawrence. The comments made and questions asked during that session were very helpful, in particular those from Lawrence and those from Yves Agid and Philipp Kellmeyer.

Cite this article: Danaher, J. and Nyholm, S. 2021. Should We Use Technology to Merge Minds? *Cambridge Quarterly of Healthcare Ethics* 30: 585–603, doi:10.1017/S0963180121000062