The Virtual as the Digital

David J. Chalmers
New York University

Abstract
I reply to seven commentaries on "The Virtual and the Real". In response to Claus Beisbart, Jesper Juul, Peter Ludlow, and Neil McDonnell and Nathan Wildman, I clarify and develop my view that virtual are digital objects, with special attention to the nature of digital objects and data structures. In response to Alyssa Ney and Eric Schwitzgebel, I clarify and defend my spatial functionalism, with special attention to the connections between space and consciousness. In response to Marc Silcox, I clarify and develop my view of the value of virtual worlds, with special attention to the case where we experience these worlds as virtual.

Keywords
Virtual reality, digital objects, spatial functionalism, transcendental idealism.

In "The Virtual and the Real", I argued for the views I called virtual realism and virtual digitalism, which together hold that (1) virtual objects are real digital objects, (2) events in virtual worlds are digital events that really take place (involving virtual properties that are really instantiated) (3) experiences in virtual reality involve non-illusory perception of a digital world, and (4) virtual experiences of a digital world can be about as valuable as non-virtual experiences of a non-digital world.

I am grateful to the commentators in this symposium for their rich and insightful comments on my paper. Of the seven commentaries, four (by Claus Beisbart, Jesper Juul, Peter Ludlow, and Neil McDonnell and Nathan Wildman) focus on issues about virtual objects, and especially on the question of whether virtual objects are digital objects. Two (by Alyssa Ney and Eric Schwitzgebel) focus on virtual properties, and especially on space in virtual worlds. One (by Marc Silcox) focuses on value in virtual worlds.

Virtual objects

Because a number of the commentaries focus on my claim that virtual objects are digital objects, let me start by setting out what this claim comes to more clearly. I presented some initial glosses on the key notions in the article, but here I will try to refine and clarify the notions and the thesis further.

In "The Virtual and the Real" (p. 314), I defined virtual objects as "the objects that are contained in virtual worlds and that we (seem to) perceive and interact with when using virtual reality". I said that paradigmatic virtual objects include "avatars (virtual bodies), virtual buildings, virtual weapons, and virtual treasures".

Digital objects were introduced as follows:

> What are virtual objects? In my view, they are digital objects, constituted by computational processes on a computer. To a first approximation, they can be regarded as data structures, which are grounded in computational processes which are themselves grounded in physical processes on one or more computers. To a second approximation, one may want to invoke more subtle relations between virtual objects and data structures, just as theorists often invoke more subtle relations between high-level nonvirtual objects (e.g. a statue) and underlying physical entities (e.g. a lump of clay). For example, in some cases, multiple data structures will be associated with a single virtual object, in which case the virtual object will be a higher-level entity constituted by these data structures. I will focus on the simple data structure view here, but much of what I say should generalize to more complex views. (Chalmers 2017: 317)

So, the first-approximation view is that virtual objects (and digital objects) are data structures. The second-approximation view is that virtual objects (and digital objects) are related to associated data structures as nonvirtual objects are related to associated physical entities, for example by constitution relations.

The key to the second-approximation virtual digitalism that I suggest is the analogy: virtual objects stand to data structures as nonvirtual objects stand to physical entities. I don't think the analogy is perfect (it is still just an approximation, after all), as there are various subtle differences between the structure of the relations in virtual and nonvirtual domains, some of which we'll see in what

follows. But I think it is on the right track.

In the target article I don't say exactly what the crucial relations are, but I give the example of a constitution relation. One clear version of virtual digitalism might rely on this relation alone, saying that virtual objects are wholly constituted by, or grounded in data structures. In fact this is not my view, for a reason I touch on in the article. Virtual worlds involve not just data structures but also human users, and many virtual objects and virtual properties are grounded in part in the minds of the human users. For example, an object is red partly in virtue of users normally experiencing it as red. Likewise, virtual money exists partly in virtue of users treating it as money. Users and their minds are not themselves grounded in data structures, or at least not in the data structures that otherwise constitute the virtual world at issue. So virtual objects need not be grounded in data structures alone.

For a third-approximation to virtual digitalism that spells out the relevant relations, we need to give a role to users and their mental states. I suggest the following: virtual objects and properties are grounded in data structures and mental objects and properties. This takes the grounding theses and adds a role for the mental in the grounding base. It covers both objects and properties, since both need to be in the grounding base. I take a mental object to be a bearer of mental states, such as a person. It is worth noting that according to the definition of virtual objects above, people (human users) count as virtual objects, since they are in virtual worlds and we perceive and interact with them. One could perhaps try to restrict usage of "virtual object" so that people do not count, but the current thesis includes them.

This refinement preserves the analogy between virtual objects and nonvirtual objects, as it is also true that nonvirtual objects and properties are grounded in physical and mental objects and properties. A possible disanalogy is that if physicalism about the mind is true, nonvirtual objects and properties will be grounded in physical objects and properties alone, while virtual objects and properties will not be grounded in data structures and their properties (at least if user's minds are independent of data structures). Of course if one is a nonphysicalist about the mind, as I am, the analogy is fine. In other work I have discussed the ways in which virtual worlds have a

structure analogous to mind-body dualism, with a realm of digital objects and a distinct realm of nondigital minds. Either way, we can perhaps flesh out the analogy by saying more explicitly: virtual objects stand to data structures as nonvirtual objects stand to physical entities (on a nonmaterialist view of the mind).

Admittedly, "digital object" is not obviously the right word for all virtual objects so construed—that is, for something grounded in data structures and mental states. Here it is probably worth distinguishing different senses of "digital object".

In the first and narrowest sense, a digital object is a *bit*: a 0 or 1 in a computational system. In the second sense, a digital object is a data structure: a computational object constituted by bits but still individuated computationally. In a third sense, a digital object includes any object wholly grounded in data structures (and perhaps other objects grounded in bits), whether or not it is itself a data structure or individuated computationally. In a fourth sense, a digital object includes any object grounded in data structures (and/or bits) and mental properties.

This is reminiscent of a familiar ambiguity for "physical object". In the narrowest sense, a physical object is a fundamental physical object (an electron, say). In an intermediate sense, a physical object is any object grounded in fundamental physical objects (a rock, say). In a broad sense, a physical object may be an object grounded in fundamental physical objects and mental states (a dollar bill, say). Philosophers sometimes use *narrowly physical* for the first sense and *broadly physical* for the second or third.

In a similar spirit, one might distinguish *narrowly digital* objects (bits and perhaps data structures) from *broadly digital* objects (objects grounded in them and perhaps in mental states). On this usage, we can understand third-approximation virtual digitalism as saying that virtual objects are broadly digital objects, where these are understood to be objects grounded in narrowly digital objects and mental states.

One might worry that the usage is too broad. For example, as it stands it counts headaches as broadly digital objects, which seems unnatural. Perhaps one could impose further conditions to restrict both what things count as virtual objects and as broadly digital objects to remove these cases. But setting these verbal issues side, at least the underlying thesis is clear enough for now.

Ludlow on the social nature of virtual objects

Ludlow's central point is that many or most virtual objects (gangs, for example) are social objects and are therefore not identical to data structures. He also goes on to argue that even data structures and other computational objects are themselves social objects, so that virtual worlds are in effect social all the way down. I am largely sympathetic with Ludlow's first point and am skeptical about the second.

Starting with the first point: Ludlow's example of a gang makes a strong case against the first-approximation view that virtual objects are identical to data structures. A gang can migrate to a nonvirtual world without the associated data structure migrating. I think that the second-approximation view and further refinements can handle these cases, though. Indeed, cases like Ludlow's involving social objects are among the central cases that motivate the refinements, because social objects are typically not identifiable with their non-social bases and centrally involve the mental.

It's a common feature of social objects (a club, a university, a city, money) that the associated underlying physical features can change greatly over time. But it remains plausible that these objects are grounded in the physical and the mental. In the same way, the digital features underlying a social virtual object can change greatly over time. But it remains plausible that the objects are grounded in the digital and the mental.

On a refined digitalist view, we can treat a virtual gang that moves from *Second Life* to *The Sims Online* to the physical world in the same way we treat a physical gang that changes locations and members and mode of operation. When the gang is in the virtual world, it presumably consists largely in a group of users and perhaps various associated digital structures (virtual buildings, uniforms, and the like). Crucial to all this forming a gang is that the users have certain attitudes, such as regarding themselves as gang members. So the gang is grounded in the digital and the mental. After partially migrating to the physical world, it might come to be grounded in a combination of the physical, the digital, and the mental. After wholly migrating to the physical world, it might be grounded in the physical and the mental with no special role for the digital.

Something similar goes for money. Physical-world currency is

grounded in a combination of the physical and the mental. Initially this might involve pieces of metal regarded a certain way. Over time, it might involve pieces of papers regarded a certain way or indeed computational structures regarded a certain way. In a similar way, within-world virtual money might be wholly grounded in the digital and the mental. If money can be used in both places, it is probably grounded in a combination of the digital, the physical, and the mental.

So nothing here threatens the key analogy. In these examples, virtual objects stand to data structures as nonvirtual objects stand to physical bases. Now, I don't have a refined positive account of the ontology of social entities and the exact relations to the physical and mental domains, but plenty of other theorists do. As far as I can tell, most of these positive accounts should apply equally well to social entities in virtual worlds, with the role of the physical replaced by the digital.

Ludlow also notes that sometimes the same object can be present in both virtual and nonvirtual worlds simultaneously. The most obvious example is a person. A hand-held controller might be another. An organization such as a financial institution might be another. These are objects in virtual worlds, so they fit the letter of our definition of virtual objects. But they may not be wholly grounded in digital and mental properties, since physical properties from the nonvirtual world may also play a role in grounding them. So these pose a problem for the third-approximation view.

One could perhaps adjust the thesis to allow physical properties in the grounding base: virtual objects and properties are grounded in digital, physical, and mental objects and properties. But this makes the thesis nearly trivial, since physicalists and dualists alike hold that everything is grounded in physical and mental objects and properties. I think it is better to exclude objects like this from the thesis by saying (as a fourth approximation) that *wholly* virtual objects and properties are grounded in digital and mental objects and properties. This restricts the thesis to objects that exist only in a virtual world, and not outside it. This may include some objects that later move outside a virtual world, as with Ludlow's gang. In this case the object will be wholly virtual and digitally/mentally grounded at the earlier time, but not later. A stronger use of "wholly virtual objects" might

exclude gangs and other objects that have users or their mental states as constituents, since these constituents are themselves nonvirtual. But I will stay with the broader use for now.

Ludlow also discusses some apparently nonsocial virtual objects, such as virtual tables. He notes that social recognition can be essential to being a virtual table, virtual sushi, and so on. I think this is true, and equally true of their physical counterparts. At least part of what it is to be a table is to be treated as a table, or to be disposed to be treated as a table, or to be designed as a table, or something like that. Two physically identical pieces of wood might differ in whether they are tables because one satisfies this condition and the other does not. On this picture, both being a table and being a virtual table will be partly grounded in the mental. Still, these will also be partly grounded in the physical and the digital respectively, and the overall picture remains a congenial one.

Ludlow's deepest concern is that data structures depend on computational states, and computational states may themselves depend on social facts. Ludlow thinks that there may be no objective fact of the matter about whether a given computational state is realized in a given physical system, if social considerations are left aside. He expresses sympathy with familiar objections from Putnam and Searle, but he is most moved by an objection from Kripke. Kripke says that what program a system is running depends on its intended function. The intended function does not depend on the physical system but on social facts including the intentions of designers and the like. If so, then computational entities may depend on social facts, so that virtual objects may be essentially social after all.

Strictly speaking, this view of computation is compatible with my virtual digitalism. Virtual objects would still be grounded in the digital and the mental—it is just that the digital would have a further grounding in the social. Still, I reject this picture. As Ludlow notes, I addressed the Putnam and Searle objections long ago, in "On Implementing a Computation" (Chalmers 1994) and "Does a Rock Implement Every Finite-State Automaton" (Chalmers 1996). As it happens I have also more recently addressed the Kripke objection, in "The Varieties of Computation: A Reply" (Chalmers 2012). My reply is roughly that which program(s) a system is intended to realize depends on its intended function, but which program(s) it is actually

realizing does not. That is simply an objective matter of the state-transitions that the system is currently disposed to make. If this is right, there need be no social component to realizing a program or the associated computational states. Of course it is true that there are some notions of information that depends on people's knowledge, as Ludlow notes with a quote from Gallistel and King. But there are also notions of computation and information that do not, and these are the notions that are most essential for my purposes.

Beisbart and Ludlow on data structures

Beisbart and Ludlow both raise interesting questions about my appeal to data structures. The basic issue is what are data structures, as I understand them? They both observe that the central notion of data structure as used by computer scientists is that of a certain abstract type (a way of organizing data) such as an array, a list, a queue, or a stack, represented in a computer program. This notion of an abstract type cannot serve my purposes, as these abstract objects seem not to have causal powers. They note that other notions tied to concrete systems may be available—Ludlow notes that philosophers often use "data structure" more broadly for something like "local semi-stable states of a computational system", while Beisbart notes that "there is something about the hardware that corresponds to the fact that the characteristics of the object (position, mass, color) take such and such values". But they both observe that the relevant notion is far from clear.

This is a fair point, so I will do my best to clarify the notion I am appealing to here. My notion is close to the one Beisbart suggests. It is the notion of a *physical data structure*, which is a *realization* of the abstract data structures specified in computer programs and the like. These physical data structures are not simply an abstract data type such as "array" or "3-element array of integers" or even a specific abstract structure such as "(3, 10, 17)". They are particular realizations of specific abstract structures in concrete computational systems, such as the realization of an array such as "(3, 10, 17)" in a particular computer.

For what it's worth, I am not sure that my use is entirely nonstandard in computer science. In computer science it is common

to distinguish between abstract data types (ADTs, e.g. queues and arrays with a certain form) and data structures. Usage of "data structure" varies: sometimes it is used more or less interchangeably with "abstract data type", sometimes it is used for more specific abstract objects, and sometimes it is used for something physical. The Wikipedia entry for "data structure" says "The ADT defines the logical form of the data type. The data structure implements the physical form of the data type". One also often finds a distinction between "logical data structures" and "physical data structures". It is true that sometimes "physical data structure" is used just for a specific way of implementing a high-level data structure by a lower-level data structure that is still abstract. But it is also often used for something tied to a physical realization. In any case, this is the notion I will concentrate on here.

It is a familiar idea that most computational entities come in abstract and concrete varieties. Abstract computational entities are mathematical objects, while concrete computational entities are present in concrete computational systems. In earlier work I have written a great deal about the realization or implementation relation between abstract computational systems (Turing machines, finite automata, and the like) and physical systems. To a first approximation, this is a matter of the concrete system having concrete states and state-transitions that mirror the formal states and state-transitions specified in the algorithm or automaton. In this case we naturally say that there is a concrete Turing machine, a concrete finite automaton, and so on.

Extending this picture: when a formal computation involves a certain data structure (say, an array of symbols on a Turing machine's tape), a concrete realization of that computation will involve a concrete realization of that data structure (say, an array of marks or of voltages realizing the symbols). These concrete realizations are physical data structures (the term "concrete data structure" might seem preferable, but this expression is already used in computer science for a sort of abstract data structure).

Let us start with the simplest data structure: a bit. Bits can be understood as abstract objects (0 or 1), but they are also realized in concrete systems, via different voltages in a transistor or something along those lines. Consider a cellular automaton such as Conway's

game of life, which involves an array of formal cells each of which can be off (0) or on (1). When the game of life is concretely realized, there will be an array of memory cells, each of which can be off (low voltage) or on (high voltage). In this way, concrete memory cells realize abstract formal cells, and concrete voltage states of the memory cells realize abstract formal states of the formal cells.

The word 'bit' has some ambiguity even when restricted to abstract data structures (setting aside uses that measure a quantity of information). For example, it can be used for a binary digit (0 or 1) in a specific abstract data structure, or it can be used for the cell whose value can be either 0 or 1 ("in the third generation this bit changes from 0 to 1"). Both of these uses have counterparts in the concrete domain. Formal cells are realized by memory cells and the like, while specific binary digits are realized by voltages and the like. It is probably most natural to say that the physical data structure realizing a bit such as 1 in a formal cell involves both the memory cell and its voltage, deriving from the cell's having a certain voltage-range. One could identify it with a property instance (the concrete bit is this cell's having high voltage) or perhaps with some other ontological category deriving from such an instance.

Likewise, an abstract array of bits such as 011001 will be realized in an array of memory cells with a corresponding array of voltages (low, high, high, low, low, high). Here the physical data structure will be something like an array of voltages, where each voltage is again a property instance such as a certain cell's having high-voltage. A more austere view might say that the data structure is simply the mereological sum of all the cells, or perhaps of all the voltages, but I think things go better with a fine-grained view where concrete objects can have mathematical structure such as the structure of an array. For example, one might see a physical data structure as consisting in the cells and voltages along with a mapping to the abstract data structure.

What about high-level data structures, such as triples of name, rank, and serial number, or the many aspects that might be encoded in a virtual object? The details will vary depending on the case, but in general it will be built into the design of a system that these high-level structures are implemented (after compilation and/or interpretation) in certain low-level structures involving bits, which are

themselves physically realized when a program is executed. To a first approximation, we could say that the physical data structure that realizes the high-level data structure is the array of concrete bits that realize the formal bits that implement the data structure. To a second approximation, one might impose more mathematical structure on the array (dividing it into three parts, and so on) so that it is structurally isomorphic to the original data structure. Then we can identify the physical data structure with this more fine-grained object.

I think this usage of "data structure" fits quite well with Ludlow's fascinating account of the creation of virtual objects in Second Life. When a new virtual object is created or "rezzed", on Ludlow's account, it is precisely a new token data structure of this sort that comes into existence.

Ludlow suggests that because the properties of such a virtual object can be changed arbitrarily, virtual objects in Second Life have no essential properties except perhaps an underlying ID number. I wondered about this. Couldn't there be a virtual statue that is essentially statue-shaped, just as many think that physical statues are essentially statue-shaped? On this view, when one changes the shape encoded in the data structure, the virtual statue is gone, though the underlying data structure will persist with a changed virtual shape (just as an underlying lump of clay might). If this is right, then even though the underlying data structures (like lumps of clay) do not have virtual shape properties essentially, the virtual objects they help to constitute might. If all virtual objects were data structures, then the statue would be identical to the data structure and this view would be unavailable (just as it would be if the statue were identical to the lump). But on a view like mine where virtual objects are constituted by data structures (just as statues are constituted by lumps), essentialism for at least some virtual objects is available.

This view addresses the worry about causal powers. As always, even if abstract objects don't have causal powers, their concrete realizations do. Concrete arrays of voltages and the like will have many effects inside a computer. Ludlow worries that it is hard for an array of values to have causal powers, especially when one thinks in terms of counterfactual dependence. But concrete arrays certainly have causal powers in this sense. Thinking counterfactually about the concrete array associated with a virtual object, it will often be the

case that if the array had contained different values, then this would have led to different outcomes, to different appearances, and so on.

In fact, Ludlow himself brings out the way that the causal powers of virtual objects (and of the associated data structures) are mediated by scripts. These scripts serve to mediate the counterfactual dependence of various effects on the array of values encoded in the data structures. This is quite congenial to my own view, and it brings out a way in which virtual causation is interestingly different from physical causation. Perhaps something analogous to scripts play a role in the causal powers of genes or of human language, but it is less clear that anything like this has a role in physics.

Ludlow worries that data structures can be eliminated through optimizing compilers and the like. (Perhaps the compiler finds a more efficient way to have the same effects with an entirely different data structure.) In this case, I would say that the abstract data structure that was encoded in the original program is not concretely realized in the post-compilation system that does the causal work. Of course there will still be other data structures in the post-compilation system. In the case where the original data structure was intended to specify a virtual object but then was replaced by one or more new data structures, the virtual object will be tied to the concrete realization of the new data structure. In the case where the data structure is eliminated entirely (it is not needed and not replaced because the virtual object was eliminated and played no role?), the corresponding virtual object may be eliminated too.

Ludlow also objects that the same data structure can appear in different ways on different systems and to different users, so they cannot be identified with virtual objects. I'm not sure I grasp the worry here. It seems to parallel a familiar phenomenon with physical objects: they can appear in different ways to different perceivers. For example, a colorful painting looks different to a color-blind person and to someone with normal vision. This is just a case of a single object appearing two different ways. Prima facie, the same goes in Ludlow's case: there is one virtual object here, appearing two different ways. The appearance of a physical object is typically quite inessential to it, and the same goes for virtual objects. When driving through fall foliage, I sometimes wear rose-tinted glasses to make the colors of the leaves appear even more striking. Modifying client-side

software to produce a new appearance for a virtual object is more or less analogous to that.

Ludlow and Beisbart both raise issues about multiple realization. Ludlow raises cases where multiple servers are involved in supporting a virtual object, perhaps storing the relevant data twice in quite different formats. In this case I'm inclined to say that the virtual object is grounded in two different physical data structures. Depending on the details of the case, perhaps there will also be a level of abstraction at which there is a single physical data structure with two different redundant realizations by more basic data structures in different locations. All of this is consistent with the version of virtual digitalism I have outlined, which is not wedded to a simple one-to-one mapping between virtual objects and data structures.

Beisbart raises doubts tied to the dynamic connection between data structures and hardware. He questions whether there are really objects that behave as I want data structures to behave, given that different hardware may realize the data at different times. In response, I think there will usually be a lowest-level of data structures (arrays of bits) that are tied directly to hardware, one formal bit per physical bit. Beisbart's worry does not arise for these. There will certainly be high-level data structures with a much looser connection both to these low-level format bits and their physical realization. Still, one can typically trace the way a high-level structure is realized at a given time, ultimately in an array of voltages and the like as before. The residual issue is that the realizers of the same high-level data structure may change over time, so we cannot simply identify the physical data structure with a specific hardware realizer.

To my eyes this issue is parallel to a familiar issue that arises with high-level physical objects. How can a hurricane be a physical object, given that it is made up of very different matter at different times? The same goes for a water wave. Certainly, one can't simply identify hurricanes with bodies of air or water waves with bodies of water. Instead, hurricanes are multiply realized across time in bodies of air. Something similar goes for the relation of concrete data structures to underlying hardware.

All this is consistent with different ontological views of what hurricanes are, and likewise of what physical data structures are. One common view is that hurricanes are four-dimensional bodies

of air across time, with different air involved at different times. By analogy a physical data structure could be a four-dimensional body of concrete bits over time, with different bits involved at different times. There are also views where hurricanes are changing three-dimensional objects, or views where they are functionally individuated by their origin, their continuity, and their effects. The same applies to physical data structure. I don't have a strong view about these matters. I think some sort of functional individuation may be best, though it is nontrivial to get the details right. There are plenty of important further questions about what data structures are, and plenty of room to clarify the notion and to clear up around the edges. But here, as usual, the digital case parallels the physical case.

It is also quite possible to state a substantive version of virtual digitalism without mentioning data structures. Such a thesis might say: wholly virtual objects and properties are grounded in narrowly digital and mental objects and properties. Here as before a narrowly digital object is a (concrete) bit, and a narrowly digital property is a value of a bit or a relation between bits. It is plausible that all wholly computational objects and properties, including data structures, are themselves grounded in narrowly digital objects and properties—in bits and their interplay. The thesis that virtual objects are grounded as above seems to capture the central interest of virtual digitalism without requiring delicate analyses of data structures.

For all I have said, it may be that the stronger thesis that replaces narrowly digital objects with data structures here is also true. At least I don't think I've yet seen a clear counterexample to it. But if it turns out that digital processes sometimes play a role in grounding virtual objects that isn't mediated by data structures, then this broader thesis will be well equipped to handle it.

In any case, physical data structures are especially useful if we are interested not so much in grounding but in the ontological question of what virtual objects are. Many virtual objects, such as those in Second Life created by "rezzing", have closely associated concrete data structures. One can either identify the virtual object with the data structure or with something in the near vicinity, such as an entity coinciding with the data structure but require to meet further conditions (as with the statue). Virtual social objects will very often be built from data structures, perhaps with further mental

conditions or components. One can expect various complexities here, but physical data structures are at least a starting point in the ontology of virtual objects.

Beisbart on simulations and the plenitude of digital objects

Beisbart rejects virtual realism on the grounds that it leads to an unreasonable proliferation of digital objects. His main argument is through a parallel with simulations, such as cosmic simulations of galaxies and the like. He argues that if all these digital objects exist in virtual worlds, a similar number of digital objects exist in simulations; but it is implausible on grounds of parsimony that all those digital objects exist in simulations; so they do not exist in virtual worlds.

I am happy to accept the first premise of Beisbart's argument. I think that at least for many of the digital objects in a virtual world, there are corresponding digital objects in a simulation. It's true that there are various differences between virtual worlds and simulations in perception, interaction, idealization, and role. These differences may make some differences to digital objects. For example, some of the mentally grounded digital objects (virtual clubs) and the like in virtual worlds may not exist in simulations where the relevant minds and mental properties are not present. Still, on my view there will certainly be many digital objects that exist in simulations, and objects like virtual galaxies will certainly be among them.

I don't accept the second premise of Beisbart's argument. I think it is independently plausible that many digital objects exist in simulations. Simulations are complex processes with many parts with many properties and relations. I would expect any such process to ground a huge number of entities. For example, if there are a billion bits (or 125 megabytes) in a simulation, then there are 2 to the billion sets of bits. If one accepts that (realized) bits are objects and one accepts universal mereological composition for objects, there will be an object corresponding to each set of bits, so there will be 2 to the billion objects on mereological grounds alone. If one is more fine-grained about individuating objects (as with the statue and the lump of clay), there will be many more still. So at least given fairly standard (if liberal) mereological principles, there will be an enormous number of digital objects.

I agree with Beisbart that the point isn't limited to simulations. In any reasonably complex computation, there will be a vast number of digital objects, for the same reason as above. Beisbart objects that this will lead to "virtual roots" in a system for extracting roots of a function which may be algorithmically isomorphic to a system in a simulation. I don't know what "virtual roots" are here and I don't see a commitment to virtual objects—since whether the digital objects in a computational system count as virtual objects (elements of a virtual world) may depend not just on the algorithm but on the way the system interacts with users. But certainly I agree that there will be plenty of digital objects in the vicinity. The same goes for the Monte Carlo process that Beisbart discusses. Whether or not the individual results of random processes count as "virtual objects", there will certainly be digital objects corresponding to them.

I hadn't thought that this view of digital objects would be especially controversial, though perhaps it would be denied by some who reject liberal mereology and the like. It's worth noting that one can get the ontology of virtual digitalism going even without liberal mereology. It is independently plausible that there are many data structures associated with a simulated object such as a simulated galaxy, and it is plausible that these data structures have concrete realizations in the computer. If so, we have digital objects for all these data structures. Perhaps someone might deny that there are concrete entities realizing data structures in computational systems, but I don't see strong reasons to do this.

Beisbart rejects all this on grounds of theoretical virtue: "*Ceteris paribus*, a theory is preferred to another one if it postulates fewer entities than the latter". It is true that Ockham's razor is often stated this way, but my impression is that this principle taken literally is very often rejected. If one astronomical theory postulates more stars than another, this is no reason at all to prefer the latter. So one needs a better formulation of Ockham's razor. One popular version is that one should prefer a theory that postulates fewer *kinds* of entities. But then, as long as we have a reasonable case that at least one physical data structure exists, the theory will be no worse off for having billions of them. Another widely discussed recent version is Schaffer's (2015) "Ockham's Laser" suggesting that we should minimize principles rather than objects. But then I see no reason why

the virtual digitalist view requires more principles than the virtual fictionalist view.

Now, Beisbart may prefer his version of Ockham's razor and may correspondingly prefer a much less liberal ontology that rejects universal composition and the like. If so, perhaps he has grounds to reject all these digital objects. Still, I think if he does this, similar considerations may lead him to reject all sorts of ordinary physical objects. After all, it is not at all clear just how and why we can get ordinary objects from a quantum wave function or whatever physics postulates as fundamental. Typically this will require quite liberal ontological principles. If we have grounds to prefer an austere ontology, we will have grounds to prefer a view where perhaps the fundamental entities exist but relatively few high-level objects exist. Perhaps Beisbart could make a case that high-level physical objects are explanatorily essential in a way that digital objects in a virtual world are not. But this is far from obvious, given both the coherence of explanations that do without high-level objects and the utility of explanations that appeal to digital objects. So I think we are at least left with our analogy in place: narrowly digital objects are to virtual objects as narrowly physical objects are to nonvirtual objects.

Beisbart makes a distinct simulation-based argument against my view along these lines: virtual worlds are analogous to simulations, so virtual objects and simulated objects should have the same status. But on the leading views of simulations, simulations are models of other systems, and the simulated objects are usually fictional objects (or sometimes real nondigital objects). If so, then virtual objects are also fictional objects (or sometimes real nondigital objects).

I think this argument goes wrong at the first premise: virtual worlds and simulations are not entirely analogous here. I agree that simulations (at least in Beisbart's sense) are essentially models and are essentially representational: they are representations of other systems, sometimes real and more often fictional. In my view, however, virtual worlds are not essentially models and are not essentially representational. Some virtual worlds are like this, as with the familiar case of videogame worlds based on fictions or historical events. However, it is a key part of my argument against virtual fictionalism that some virtual worlds are not representational or fictional at all. Many workaday virtual worlds in the future will simply be spaces for

work or interaction without any special representational function. It is implausible to say that the virtual objects in these worlds are fictional objects.

This argument against virtual fictionalism does not apply to fictionalism about simulated objects. Insofar as all simulations are representational, it is much more plausible that all simulated objects are represented objects, which will typically be fictional. Of course I think that the relevant digital objects will still be present in simulated worlds. It's just that it's not especially plausible that a simulated galaxy is a digital object. The simulated galaxy is a (fictional) physical galaxy—that's the whole point of the simulation. We might give the corresponding digital object a different name, such as a "simulation galaxy" or a "galaxy representation" or some such. Anyway, precisely because of their essential representational function, I think simulations need to be treated quite differently from virtual worlds.

McDonnell and Wildman on strong and weak virtual digitalism

McDonnell and Wildman say that my view is ambiguous between strong virtual digitalism, according to which virtual objects are digital objects, and weak virtual digitalism, according to which virtual objects are dependent on digital objects. In fact, I say that virtual objects are identical to digital objects, and I canvas strong and weak views of the relationship between digital objects and data structures (preferring on second approximation the view that digital objects are grounded in data structures). It may be that M&W are using "digital objects" more narrowly than me to just encompass data structures and not higher-level entities grounded in them. Using the expression that way, my first-approximation view is their strong digitalism and my second-approximation view is closer to their weak digitalism.

Moving to M&W's substantive arguments, they first object to my two main arguments for virtual digitalism. They are here questioning mainly what they call strong digitalism, which I reject at least when understood their way, but as far as I can tell their arguments remain interesting when understood my way and applied to the view I accept. As M&W eventually note when discussing my different use

of the terms, different usages just move the bump in the rug, so I will construe all of their arguments as arguments against my actual view to see if they have any force.

My argument from causal powers says roughly: virtual objects have certain causal powers, digital objects are the objects with the relevant causal powers, so they are virtual objects. M&W reject this on the grounds that users have causal powers too. I would reply that users don't have the relevant causal powers—those that we want to say virtual objects have. M&W say this restricts the scope of the argument dramatically, but as far as I can tell it applies to any wholly virtual object that interacts with other virtual objects or that appears a certain way—which is just about all of them. They also worry that the relevant causal powers include effects on virtual objects and it's far from obvious that digital objects have those effects. Perhaps they have effects on other digital objects, but it would be question-begging to assume that these are virtual objects. That's a fair point—but there are at least pretty clear counterfactual relations between digital objects and virtual objects downstream, so at least on a counterfactual analysis of causation there is a prima facie case for a causal connection here.

My argument from perception says roughly: we perceive (only) virtual objects in VR, and we perceive digital objects because they cause our perceptual experiences, so virtual objects are digital objects. M&W reject this on the grounds that we also perceive nondigital objects such as other people. That's also a fair point. To fix the argument we can just restrict it to wholly virtual objects, as before. Roughly: when I perceive (only) a wholly virtual object, I perceive a digital object because it causes my experience, so wholly virtual objects are digital objects. This avoids M&W's worry.

M&W also raise cases of crossplay, where the same virtual object (a virtual frisbee, say) is realized by two physical data structures on different servers. I handle this the same way I handled a similar discussion from Ludlow: this is a case where a single virtual object is realized jointly by two data structures. This is a problem for the view that virtual objects are identical to data structures, but no problem for the view that virtual objects are grounded in data structures. It is only a problem for the "strong digitalism" thesis that virtual objects are digital objects if one further identifies digital objects with data

structures, as M&W do but I do not.

M&W go on to address their "weak digitalism": virtual objects are grounded in digital objects, i.e. in data structures. The main objection on this view is that data structures will do all the causal work, and virtual objects won't be in the causal loop. The argument is a classic exclusion argument, analogous to arguments on which (say) narrowly physical objects and properties do all the causal work, so high-level entities such as biological objects and properties will be out of the causal loop. Like most philosophers, I think these arguments fail. In my view they fail especially when there is a grounding connection between the lower-level and higher-level objects and properties in question. If biological properties are grounded in narrowly physical properties, then causal work by the narrowly physical does not exclude causal work by the biological. Similarly, if virtual objects and properties are grounded in digital objects and properties, then causal work by the digital does not exclude causal work by the virtual.

M&W consider this sort of line but say it only works because the biological and narrowly physical descriptions are redescriptions of the same phenomenon. And they say that we cannot take this like in the digital-virtual domain since this would lead to strong digitalism where virtual objects are identical to digital objects. Now, I don't know exactly what it means to be a redescription of the same phenomenon, but if this entails object identity, it very plausibly doesn't apply in the biological case either—since biological entities very plausibly aren't identical to narrowly physical entities. And if it doesn't entail object identity, there's no longer reason to think the view collapses into strong digitalism. In any case I'd prefer to use grounding rather than "redescription" as a criterion for when exclusion arguments fail. If this is right, then grounding-based versions of virtual digitalism are unthreatened.

M&W go on to advocate "virtual walt-fictionalism": a version of virtual fictionalism tied to Kendall Walton's version of fictionalism where fictions centrally involve a structured practice of make-believe based on props that engage one's imagination. I think that this is a coherent view and it certainly covers some virtual worlds. Most obviously, it covers the virtual worlds of role-playing video games, which essentially involve fictions and a certain sort of pretense. It

may also capture some users' attitude toward non-videogame virtual worlds for some users, when these users experience them via attitudes of pretense that they are in a physical world with the same features. But I don't think that make-believe is essential to the experience of virtual reality.

M&W don't address my main discussion of fictionalism in section 6, but some points made there apply directly to walt-fictionalism. For sophisticated users of VR in many VR environments (say, a non-videogame world used for work or communication), fictional interpretations of the world need not be at all salient, and pretense and make-believe may play no role. One is in a virtual world and one experiences oneself as in a virtual world. One need not experience oneself as in a physical world, and there is no pretense. One simply interacts with other people or with virtual objects, in much the same mode as one interacts with other people or physical objects in the non-virtual world. One's imagination need not be engaged in any special way beyond the ways it is engaged in non-virtual interactions. Perhaps in principle we could eventually test this claim empirically, with good enough empirical criteria for identifying acts of imagination and make-believe. In any case, if I am right, then walt-fictionalism is not true of virtual worlds in general. For a general view of virtual worlds we will do better to endorse virtual digitalism.

Juul on the half-reality of virtual objects

Juul thinks that virtual objects are at best half-real and half-fictional. His reason is that virtual objects are never full-blown counterparts of physical objects with all their complexity and detail. Instead they are heavily simplified objects. Juul suggests that this makes virtual objects incomplete and partly fictional.

Of course Juul is right that current virtual objects do not have anything like the complexity and functionality of ordinary physical objects. Nothing in my article depended on assuming otherwise. It's true that in discussing value in virtual worlds, I talked about the future possibility of high-quality virtual worlds where objects are as rich as physical objects. Juul suggests in the title of section 2 that these full-blown virtual objects are "impossible", but then goes on to say that because of all the physics-based simulation that would be

required they require "computational powers well beyond existing technology". I agree, but this is a far cry from impossibility.

In any case I am happy to focus on ordinary virtual objects that exist today. I intended my main theses to apply to these objects. This goes especially for the first three theses about virtual objects, virtual events, and perception in virtual worlds. The fourth thesis about value involved a "can" claim about what is possible, and I concede that the simplifications involved in current virtual objects may involve some diminution of value. But I think these objects remain fully real.

As far as I can tell, Juul's main points would apply equally well to mechanical replicas of ordinary entities as to virtual versions of them. A robot kitten is a vastly simplified version of a kitten. A toy car is a very simple version of a car. Nevertheless, they are perfectly real physical objects, and their simplicity does not detract from their reality as objects.

What is true is that a robot kitten is not a real kitten—not even close. And a toy car is not a real car. But this is a different issue. My sense is that for some theorists in this area (perhaps including Juul), the "real X" issue is the main issue at play in discussions of reality vs fictionality. If a virtual kitten is not a real kitten, it is to that extent at least partly unreal and fictional. But if that is the main issue, I have long ago conceded it. In the target article, I am clear that I think virtual kittens are not real kittens and so on. I don't think one needs to focus on simplified counterparts to make that point. Even a full-blown simulation of a kitten (if possible) would plausibly not be a real kitten, just as a full-blown simulation of a hurricane is not a hurricane.

It is true that simplified virtual simulations of ordinary physical objects might be seen as representations of those objects. Because of the simplifications, they are *misrepresentations*: they do not to justice to the real thing. That might yield another sense in which these virtual objects are fictional. Still, as I discuss in the target article, virtual objects need not always be representations of physical objects. Some are simply virtual objects in their own right, not representing anything else.

An analogy here is a mechanical robot device such as a Roomba, which (unlike a robot kitten) is not intended to represent or simulate any physical object. A Roomba is completely real. It is in no sense

fictional, because it is not trying to simulate something outside itself. The same goes for many virtual objects—the stylized virtual objects in a game such as Tetris, for example. These are not really trying to represent a physical domain of tiles falling from the sky. They are just virtual objects in their own right. As such, they are perfectly real virtual objects.

I am not sure that Juul and I really disagree substantively here. We may just be using words such as "real" or "fictional" in different ways. Juul even allows that virtual objects are "real in a different special sense, where we as users understand that they are specific, stylized versions of their non-virtual counterparts." I may not call that a different sense of "real", but I agree that virtual objects typically have this property. I am still not sure how Juul would classify virtual objects that are not counterparts of physical objects. Perhaps they are still "fictional" because their capacities fall short of those of genuine physical objects. On the other hand, Juul himself notes that after a while we don't expect virtual objects to be as complex as physical objects, and we take them on their own terms. At that point I think misrepresentation and fiction largely disappear from the scene, and we are dealing with objects that are fully real.

Ney on phenomenal functionalism about color and space

Ney focuses on the broadly functionalist view that I advocate for our reference to color properties and spatial properties, both inside and outside virtual reality. On this view, color properties and spatial properties are picked out in virtue of the role they play, with a special emphasis on a role in causing color and space experiences. She suggests that this "phenomenal functionalism" might have some plausibility for colors but that it is quite implausible for space. Color properties are tied much more closely to their appearances than spatial properties are.

Ney's central argument starts from the phenomenon of color metamerism, where two objects with quite different surface properties look to be the same color under normal conditions. Ney says that in this case, we always say the objects have the same color. So that is consistent with the claim that something is red if and only if it looks red under normal conditions. In the case of space, however, there

are various spatial illusions where two objects look to be similar spatially under normal conditions, but they nevertheless have quite different spatial properties. One can get these cases going with illusions where objects are made to look taller than they are, or to look hollow when they are not. This is inconsistent with the claim that something is hollow if and only if it looks hollow under normal conditions, or that something is tall iff it looks tall under normal conditions. If so, phenomenal functionalism is false for spatial properties.

One could perhaps quibble with some of Ney's claims about illusions. There are certainly some cases of color illusions under normal conditions. For example, in versions of the well-known checkerboard illusion two squares look to have the same grey color when they have quite different colors (perhaps one is light and one is dark, objectively). Now, Ney might reply that in this case conditions are abnormal: the square only looks grey because of some unusual context and apparent lighting. But if we are allowed to make this sort of move, we can make the same move in Ney's spatial cases. The convex object normally looks convex, and only looks hollow due to an unusual context. The person who looks tall normally looks shorter—they just look tall here due to unusual conditions. Still, one can certainly find cases of spatial illusions under normal conditions. Green and Rabin (2019) point out a number of these.

My main response to Ney's arguments is that it only refutes an especially strong version of spatial functionalism: one that in *Constructing the World* I called *simple spatial functionalism*. This version says that for every spatial property S, something has S iff it looks S under normal conditions. In effect, this says that the referent of every spatial concept or spatial expression is fixed wholly as the property that causes the associated type of spatial experience. I have occasionally put this view forward as a first approximation to spatial functionalism, but it is certainly not my considered view.[1]

---

[1] Green and Rabin 2019 give an extended argument against my spatial functionalism that turns in significant part on attributing simple spatial functionalism and an associated "no lifelong spatial illusions" thesis to me. I certainly reject both, as discussed in the main text. In Chalmers 2019a (their main target) I argued that certain special lifelong spatial illusions involving systematic global inversions of orientation, size, and relative size are impossible. I did not endorse the broader claim that no lifelong spatial illusions are possible and take it to be

It's true that in the target article my exposition of color and spatial functionalism was extremely brief, and there are a few remarks that may suggest simple spatial functionalism. For example, I say "That is, to a first approximation, an object is one meter tall when it normally causes experiences of being one meter tall… And so on." But even here, I say it is a first approximation and that I am ignoring many subtleties. Furthermore, I go on to invoke a separate defining causal role for spatial properties, tied to their role in dynamic interaction and motion: "Distance is what there's no action at" and "what there's no motion at". (I likewise say "Virtual objects exist in their own virtual space, in virtue of their effects on each other *and* on our experiences."). This already goes well beyond simple spatial functionalism. The more complex picture of reference to spatial properties opens up the possibility and likelihood of spatial illusions where objects have spatial properties other than those they seem to have.

Another difference with simple spatial functionalism that I have discussed in other work is that the reference of spatial concepts, spatial expressions, and spatial experiences is fixed *holistically*. That is, the reference of the whole manifold of spatial concepts (and so on) is fixed at once, in a way that respects various internal relations between elements of the manifold. Part of this involves respecting continuity relations and the like (continuous varying spatial experiences should be assigned continuously varying referents) while part of it involves respecting equivalence relations (e.g. distances that look the same should be assigned the same referent, *ceteris paribus*).

Another crucial holistic aspect of both color and spatial reference-fixing is *compositionality*. A certain purple-ish experience (where an X-ish experience is an experience as of X) is plausibly compositional, involving a certain amount of blue-ishness and a certain amount of reddishness. Its content might be "that object is 0.5B and 0.3R"— where B is the property that normally causes blue-ish experiences, and R the property that normally causes reddish experiences. Because of compositionality, it is possible for the purple-ish experience

---

obviously false. The illusions that Green and Rabin discuss raise some interesting questions but all can be analyzed using the holistic and compositional framework outlined in the main text. These illusions pose no real threat to a more sophisticated spatial functionalism, as Green and Rabin go on to allow.

to be illusory even in normal conditions. Suppose experiences of this kind are always caused by objects that have 0.2B and 0.8R, for some reason tied to the interaction of these two specific shades and that arises in no other cases. Then these experiences will be illusory. If there was no compositionality the experience would represent its normal cause, without illusion. But compositionality allows normal illusions. In effect, reference is fixed in a relatively simple way for the primitive dimensions of color experience, so it will be relatively hard to get normal illusions for the primitive dimensions. But it will be straightforward to obtain normal illusions for compositional experiences, in cases like this when a normal cause for the compositional experience involves abnormal causes for its parts.

All this applies much more to the case of spatial experiences, which have a rich compositional structure going far beyond the structure of color experiences. Visible objects have shapes, sizes, and locations; they are made up of lines and points and angles; perceived locations in space involve distance from one and a relative angle to various axes, and/or perhaps a location on depth, left-right, and up-down axes. Perhaps the content of the individual components is fixed in something close to the simple way. In "Three Puzzles about Spatial Experience", I argued that there are limits on the sorts of normal illusions one can experience with elements such as left-right orientation and size. But it is easy to have normal illusions involving experiences of complex spatial properties such as hollowness and location.

So simple spatial functionalism needs to be modified into a number of ways. It needs to be modified to allow for compositionality and to allow for other forms of holism, and it needs to allow nonphenomenal roles as well as phenomenal roles to be involved in reference-fixing. With these modifications in hand, one will expect the connection between spatial properties and their appearances to be relatively weak. The cases Ney mentions will then be no threat. There may still be some connections of the sort mentioned above, and perhaps Ney thinks even these connections are too strong. If so, the ball is in her court to deliver illusions or other phenomena that call even these weak connections into question.

I also note that nothing in my spatial functionalism entails that space does not have mind-independent reality. The spatial functionalism is a story about how reference to space is fixed, not about the

nature of space itself. A mind-independent space is very well-suited to playing the causal roles in question, so the picture is quite congenial to the mind-independence of space.

I would expect that everything I have said applies in principle to virtual space. There can in principle be normal spatial illusions in virtual worlds for reasons tied to holism, compositionality, and nonphenomenal roles. And although spatial experience plays a role in fixing reference to virtual space, virtual space may well be mind-independent.

That said, I am not sure I agree with Ney's suggestion there has to be some special similarity between physical space and virtual space. On my view these spaces will be similar in their causal roles and perhaps their abstract structures, but the properties that play the roles may be quite different in the two cases. For a spatial functionalist, role similarity is similarity enough.

Ney also suggests that facts about color and shape properties in virtual worlds can be fixed in a way quite different from in ordinary reality—roughly, by a sort of stipulation. She considers an avatar in *Gun Club* which appears to have a round belly, but which is accompanied by a dialogue saying that the character is actually skinny but in disguise. Ney says that in this case the avatar is virtually skinny, despite the apparent shape, and that no view on which virtual shapes work like nonvirtual shapes can account for this. The same goes for a case where a virtual character looks tanned but is stipulated to be pale. Ney thinks the character is pale, and that views of virtual color like mine cannot account for this.

I don't have clear intuitions about whether the character is fat or skinny, or pale or tanned, in this case. I think intuitions can go either way. We need to know more about the virtual world. If the virtual world is clearly intended as a fiction, with an associated narrative, then the stipulation is plausibly part of the narrative and helps constitute facts in that world. On the other hand, if the virtual world is a workaday world used for communication or employment, I'm not sure that stipulations like this make any difference. People are not interested in these worlds as representing a distinct reality. They're interested in it as its own reality.

So my first pass is that Ney is right for virtual worlds qua fictions or representations, but not for virtual worlds qua realities in their

own right. Fixing the facts about a fiction can involve all sorts of methods, including Ney's stipulations. Fixing the facts about a directly experienced reality is a different matter. I think the stipulation can only take hold if the participants have an attitude of make-believe—which, as I said earlier, holds of some videogames but certainly not of virtual worlds in general.

As before I think that in fiction-involving videogames and the like, there is both a fictional world and a digital world present, where the latter represents the former. Likewise, there is both a fictional character and a digital avatar that represents it. I am inclined to say the fictional avatar is physical, pale and skinny (though perhaps there is also another associated fiction where the avatar is physical, tanned, and round). But the digital avatar is virtual, tanned, and round—that is, virtually tanned and virtually round. And that is what my account of spatial and color properties predicts.

Schwitzgebel on transcendental idealism

Schwitzgebel uses ideas about virtual worlds, and especially the idea that we might be living in a simulation, to articulate a way in which Kant's transcendental idealism might be true. Schwitzgebel's version of transcendental idealism (call this STI) has two components, what I'll call *humility* and *spatial mind-dependence*. Humility says that the fundamental nature of things as they are in themselves, independently of us, is unknowable to us. Spatial mind-dependence says that spatial properties depend on our minds (external things, independently of us, do not have spatial properties).

As Schwitzgebel notes, I am sympathetic with the broad project. I've said elsewhere that I take the simulation idea as an illustration not of Cartesian skepticism but of Kantian humility. At least, I take it to be an illustration of the sort of humility that Rae Langton calls Kantian humility, which is closely related to the sort that David Lewis calls Ramseyan humility. The rough idea of this sort of humility is that sciences gives us knowledge of certain structural features of the world (which might be equated with phenomenal features) but not of their underlying realizers (which might be equated with noumenal features). The simulation hypothesis is one in which the realizers involve processes inside a computer in the next level up.

They are noumenal at least for us, since (at least in a perfect simulation that lacks red pills and the like) we can't discover that our world is realized this way. But we can know our phenomenal world, which is perfectly real. I don't really think this scenario is really very close to Kant's view of reality. For a start, the next-level-up world of the simulators does not seem to be deeply noumenal (it is perfectly perceivable by beings in that world), and it has a structure and perhaps realizers of its own. But still there is a Kantian flavor.

The spatial component of Schwitzgebel's transcendental idealism is at least closely related to phenomenal spatial functionalism, which (in a version) says that space is picked out as whatever causes our spatial experience. On this view, space counts as spatial because of the way we experience it. Does this mean that my view is a version of transcendental idealism? I think it depends on how "spatiality depends on the mind" is defined. On my view, which is a sort of *realizer* functionalism about space, spatial experience plays a central role in fixing reference to space, but space itself may well be mind-independent. One way to put it is that there is a conceptual dependence of space on the mind but not a metaphysical dependence. If Schwitzgebel's transcendental idealism requires the first sort of dependence, then I may be a transcendental idealist in this sense. If it requires the second sort (as I think it does), I am not. On the other hand, a closely related *role* functionalism about space, where spatial properties are identified with something like powers to cause spatial experience, will have a metaphysical dependence of space on the mind and so will be a sort of transcendental idealism in this sense.

Interestingly, on two occasions when I gave a paper about this sort of spatial functionalism, Kant scholars in the audience (Anja Jauernig and Allen Wood) told me that they thought the view was a sort of transcendental idealism ("space is the form of intuition"). I don't have enough expertise on Kant to know whether that's right, and my view is different from Kant's in many other ways, but it's still an interesting thought. It's worth noting that spatial functionalism alone is not committed to the humility thesis with an unknown intrinsic nature to space, but the two views are at least consistent.

Schwitzgebel thinks the ordinary picture of a simulated reality does not really deliver transcendental idealism. His reason is that the simulation will itself be grounded in a physical world with space,

and that space will be independent of our minds, contrary to the required claim about space. Now, I'm not sure that the space-like structure that governs the simulating world is itself *space*. "Space" plausibly refers to the space-like structure in our own world, so perhaps the structure in the containing world (which after all may have very different laws) is at best spacelike, or perhaps "a space". If so, then if we are in a simulation, the space of the next world up may be mind-independent, but space itself (that is, the space of our world) may be mind-dependent, at least if the right sort of phenomenal spatial functionalism is correct. So I think even an ordinary simulated reality plus spatial functionalism may deliver a version of Schwitzgebel's transcendental idealism.

Schwitzgebel focuses on a stronger view that he takes to be a more genuine transcendental idealism. This is a view where the universe is grounded in a simulation taking place in one being's immaterial mind. Schwitzgebel says that this world is fundamentally nonspatial, and the only spatial properties in the world depend on our minds. My view as before is that we don't need to ground all reality in the mental to vindicate Schwitzgebel's transcendental idealism, but this is one way to get there.

In some respects, this view looks less like Kant's transcendental idealism and more like Schelling's objective idealism. Schelling's view replaced Kant's unknowable noumenal realm with a mental realm. Schwitzgebel does the same. The Schelling/Schwitzgebel view is clearly a version of metaphysical idealism, where Kant's view is usually taken to be at most a version of epistemological idealism. On the other hand, Kant's epistemological idealism is perhaps consistent with a metaphysical idealism where the unknowable noumenal world turns out to be mental.

I agree with Schwitzgebel that his sort of idealism is a real possibility worth discussing. I discuss it under the label of "cosmic idealism" in a recent article on idealism and the mind-body problem (Chalmers 2019b). I am not sure that cosmic idealism can solve the mind-body problem. It faces a version of the familiar "decombination problem" for cosmopsychism, and needs to answer the question of how our minds could be grounded in the cosmic mind. But still it is a view worth attending to.

Schwitzgebel worries that on some "structuralist" or spatial

functionalist views, a cosmic-mind simulation will contain space independently of our minds. The worry is less pressing for phenomenal spatial functionalism, which makes space depend on our minds—but as before this dependence may only be conceptual dependence, not metaphysical dependence, so a lot turns on which of these matter for Schwitzgebel's transcendental idealism. Schwitzgebel's footnote 14 on phenomenal spatial functionalism suggests it is the latter, in which case his transcendental idealism may be threatened. On this picture spatial properties will be properties internal to the cosmic mind, and will depend metaphysically on the cosmic mind and not on ours.

Schwitzgebel's transcendental idealism is more clearly threatened by a nonphenomenal spatial functionalism, where space is picked out as what plays a certain dynamic role. On a nonphenomenal spatial functionalism, space will still exist independently of us in the immaterial mind. Still, even on this sort of view, space will not exist at the *fundamental* level of the mind-independent world. Instead it will be a sort of high-level feature of that world, albeit one that does not depend on us. Perhaps that can be considered a small step away from the strongest form of realism about space and at least a step in a Kantian direction.

Silcox on the value of virtual worlds

Silcox addresses my suggestion that virtual worlds can be as valuable as nonvirtual worlds. He says there is a way in which my virtual digitalism undermines this claim. He focuses especially on my suggestion that the "phenomenology of virtuality", in which we experience objects in a virtual world as virtual, helps us to avoid illusions and thereby makes experiences more valuable. Silcox thinks there are numerous situations in which we obtain more value when we regard virtual objects as *physical* than as virtual. If so, the phenomenology of virtuality is at best a double-edged sword.

Much of Silcox's discussion concerns videogames. He says that many videogame players obtain more value when they experience the game environment as physical and not as virtual. Virtual rocks are more interesting when one sees them as physical rocks. A quest is more engaging when it is experienced as a physical quest. If we

experience these things in a "sophisticated" way as virtual, we lose value. The zombies in horror games will be less horrifying if we experience them as merely virtual.

I didn't make claims about the value of videogames in my article and I have no real expertise here. I am prepared to allow that what Silcox says may be true for at least some videogame players. I suspect that there are others for whom it is not, and for whom experiencing videogame objects as virtual does not detract from their enjoyment. But as always, I think videogames are a special situation. Videogames typically involve a fiction that is central to experiencing them, and very often that fiction may involve the fiction of a physical world. In these cases, it will not be surprising if experiencing virtual objects as physical is essential to properly appreciating the fiction and to experiencing the game in the way it was intended.

Silcox also suggests that on a Suits-inspired account of gameplay, the complexity of virtual objects may not be essential to appreciating them. One expects them to have only the features required for game-specific purposes, so they can be much simpler than nonvirtual objects. Now, I will not be surprised if future games with increased complexity and functionality of virtual objects will involve enhanced appreciation even as games. I suspect that if VR that nearly replicates a full-scale nonvirtual world is ever possible, recreational uses of these worlds for gaming may be quite popular. Perhaps for example one could experience a gripping and engaging full-scale virtual simulation of walking a tightrope or climbing Everest, without the associated danger. But in any case, I am happy to allow that where many current videogames are concerned, complexity is inessential to their enjoyment.

Silcox himself allows that once we move away from videogames to VR used for other purposes (teleconferencing, experimenting, speed dating), perceiving virtual objects as virtual may well have positive effects. So I am not sure we have a deep disagreement here.

Later in his article, Silcox makes a similar point about Nozick's experience machine. He suggests that part of the appeal of the experience machine is that one does not know one is in it. One erases one's memory of stepping into an experience machine in order to have a more immersive experience. If a subject was "sophisticated" and experienced the experience machine as an experience machine,

the experience would be less valuable.

Silcox's suggestion here is echoed in Cypher's famous remark in *The Matrix*: "Ignorance is bliss". Perhaps Silcox is right that in cases like these, ignorance enhances enjoyment. But enjoyment is not the same as value—at least, that was the main point of Nozick's experience machine discussion. Silcox suggests that his view about the value of ignorance is implicit in Nozick, but I don't think that is quite right. Nozick will presumably allow that ignorance can enhance enjoyment, but the tenor of his discussion suggests that he does not think it would enhance value.

As for me: In discussing the Experience Machine in the target article, my main point was that Nozick's central reasons not to plug into the Experience Machine, which turn on it being illusory and preprogrammed, do not apply to VR. I didn't take a stand on whether these reasons are good reasons. For what it's worth, I don't think I would step into a long-term Experience Machine myself, in large part because of the memory erasure (and also the preprogramming). I'd much rather keep my memories and have clear-eyed knowledge of where I am. But perhaps there are others who would see things differently. If so, more power to them. Different things are valuable for different people.

In any case, even if Silcox is right about all of these things, I don't think anything calls into question my main thesis that virtual worlds can be as valuable as non-virtual worlds. Silcox's points concerns the relative value of two ways of inhabiting virtual worlds (experiencing them as virtual or as physical), and not the value of these compared to non-virtual worlds. I think that experiencing them as virtual is a positive, in part because knowledge and authenticity is a positive. Silcox suggests that reality and authenticity are over-rated. That is a long discussion. But whichever view is correct, the answer to this question is entirely compatible with virtual worlds being about as valuable as nonvirtual worlds.

David J. Chalmers
Department of Philosophy, New York University
5 Washington Place
New York NY 10003, USA
chalmers@nyu.ed

References

Beisbart, C. 2019. Virtual realism: really realism or only virtually so? A comment on D. J. Chalmers's Petrus Hispanus Lectures. *Disputatio* 11(55): 297–331.

Chalmers, D. J. 1994. On implementing a computation. *Minds and Machines* 4(4): 391–402.

Chalmers, D. J. 1996. Does a rock implement every finite-state automaton? *Synthese* 108(3): 309–33.

Chalmers, D. J. 2012. The varieties of computation: A reply. *Journal of Cognitive Science* (3): 211–48.

Chalmers, D. J. 2017. The virtual and the real. *Disputatio* 9(46): 309–52.

Chalmers, D. J. 2019a. Three puzzles about spatial experience. In *Blockheads! Essays on Ned Block's Philosophy of Mind and Consciousness*, ed. by A. Pautz and D. Stoljar. MIT Press.

Chalmers, D. J. 2019b. Idealism and the mind-body problem. In *The Routledge Handbook of Panpsychism*, ed. by William Seager. New York: Routledge.

Green, E. J.; and Rabin, G.O. 2019. Use your illusion: Spatial functionalism, vision science, and the case against global skepticism. *Analytic Philosophy*.

Juul, J. 2019. Virtual reality: fictional all the way down. *Disputatio* 11(55): 333–43.

Langton, R. 1998. *Kantian Humility: Our Ignorance of Things in Themselves*. Oxford University Press.

Lewis, D. 2009. Ramseyan humility. In *Conceptual Analysis and Philosophical Naturalism*, ed. by D. Braddon-Mitchell and Robert Nola. MIT Press.

Ludlow, P. 2019. The social furniture of virtual worlds. *Disputatio* 11(55): 345–69.

McDonnell, N.; and Wildman, N. 2019. Virtual reality: digital or fictional? *Disputatio* 11(55): 371–97.

Ney, A. 2019. On phenomenal functionalism about the properties of virtual and non-virtual objects. *Disputatio* 11(55): 399–410.

Nozick, N. 1974. *Anarchy, State, and Utopia*. Basic Books.

Schaffer, J. 2015. What not to multiply without necessity. *Australasian Journal of Philosophy* 93: 644–64.

Schwitzgebel, E. 2019. Kant meets cyberpunk. *Disputatio* 11(55): 411–35.

Silcox, M. 2019. The transition into virtual reality. *Disputatio* 11(55): 437–51.