An Ontology of Ideas

ABSTRACT: *Philosophers often talk about and engage with ideas. Scientists, artists, and historians do, too. But what is an idea? In this paper, we first motivate the desire for an ontology of ideas before discussing what conditions a candidate ontology would have to satisfy to be minimally adequate. We then offer our own account of the ontology of ideas and consider various strategies for specifying the underlying metaphysics of the account. We conclude with a discussion of potential future work to be done on the ontology of ideas.*

KEYWORDS: ideas, mental states, ontology, philosophy of art

1. Introduction

Consider the idea that gravity is not a force, but the curvature of space. Albert Einstein generated this idea. He wrote it down and shared it with his colleagues. Now the idea is shared by physicists around the world. Since its origination, it has proven to be quite influential. But, an inquisitive ontologist might ask, what sort of entity did Einstein bring into existence through this generative activity? That is, what *is* an idea?

In this paper, we offer an ontology of ideas, with a focus on developing such an ontology against a backdrop of discussions in contemporary analytic metaphysics and philosophy of mind. The account on offer here is a more fully developed version of the account discussed in Cray (2014) and originally introduced in Everett and Schroeder (2015). (For a thorough discussion of the ontology of ideas in pre-Kantian modern philosophy, see Hight [2008].) We favor this ontology because it captures widely held intuitions about ideas, does so in a simple and parsimonious manner, and enjoys theoretical fecundity. We start by motivating the desire for an ontology of ideas. Next, we discuss various criteria that any adequate ontology of ideas should satisfy. We then present our own ontology of ideas that, we argue, satisfies these criteria and go on to explore various metaphysical frameworks within which we might situate this ontology. Our ultimate goal is modest: as philosophers,

we—like almost everyone else—spend a lot of time and energy talking about ideas, and we aim here to start a conversation about what such entities, so common and central to our practices, *are*.

2. The Roles of Ideas

Ideas matter across a range of contexts. Understanding the origins and development of ideas is integral to the study of intellectual history: we can inquire into the influence ideas have had on various cultures and ask how ideas have influenced other ideas over time.

An understanding of the origins and development of ideas is also relevant to issues of intellectual property. It is widely held that I hold some sort of dominion over ideas that are *mine* and hold less sway over ideas that are not mine. (For relevant discussion that to various degrees anticipates the one developed in this paper, see Trerise 2010.) We can debate, as Leibniz and Newton famously did, whether two people independently invented their ideas or whether one person stole an idea from another, as in cases of plagiarism.

More generally, we often praise or condemn people for ideas they have invented or helped spread. Consider, for example, the celebration by some (and demonization by others) of Charles Darwin for his ideas about evolution through natural selection, the castigation of George Lucas for his ideas (and perhaps now of Disney for its ideas) about how best to carry on the *Star Wars* franchise, and debates about missionaries spreading religious ideas. We might censure those who censor, for blocking the spreading of ideas, and applaud those who encourage an environment friendly to the production of new ideas.

Of course, a fully worked-out ontology of ideas is not *necessary* for any of the above to make sense. But cultivating an understanding of what ideas are will certainly help us better understand the particulars of their origins and development, which will illuminate our understanding of the aforementioned topics.

Ideas have also been invoked in certain specialized areas of philosophy. Consider the ontology of art—in particular, late modern art. In a sentiment echoed by many, Joseph Kosuth (1969: 166) has said that in such art, 'the "art idea"... and the art are the same'. Sol LeWitt (1969: 107) writes that 'ideas alone can be works of art' and that 'all ideas are art if they are concerned with art and fall within the conventions of art'. (For more discussion on the relationship between late modern art and ideas, see Goldie and Schellekens [2007, 2010], and Cray [2014].)

Nearby, in the ontology of music, Renée Cox (1986) and Benedetto Croce (1909) have advocated a similar view, as has R. G. Collingwood (1938: 142), with the latter suggesting that such works are ideas (or some similar mental entities) when he said that '[t]he work of art is not the collection of noises that result when the tune is played, but the tune in the composer's head'. (For criticism of such views, see Dodd [2007: 26–30] and Wollheim [1980: 34–43]; for alternative interpretations of Collingwood, see Ridley [1997] and Davies [2008].) Elsewhere, in the philosophy of fiction, Everett and Schroeder (2015) suggest that fictional characters are ideas used in telling stories. A clear understanding of the ontology of ideas will help in

evaluating and making such claims more precise, claims that themselves might fare better or worse depending on the ontology we ultimately settle on.

Ideas might also be suited to other roles in philosophical theorizing. Perhaps we would find it helpful to take *scientific theories* or *hypotheses*—such as Darwin's theory of natural selection—to be *ideas about how (some aspect of) the natural world works*. If so, as David Hull (1978: 356) suggested once, then work in the ontology of ideas has some import—perhaps small, but not insignificant—for the philosophy of science. Perhaps *historical narratives*—such as competing views about what happened during the course of any particular war—are *ideas about how (some part of) the world was*. If so, then work in the ontology of ideas has utility within historiography and the philosophy of history. And insofar as we might talk about *ideas for how we should act*, *ideas for how to derive a certain mathematical conclusion*, *ideas for making things a certain way*, and so on, work in the ontology of ideas connects with other areas of philosophy as well.

Of course, these claims are all controversial, but at least this much is clear: whatever their scope and however we apply them, ideas are clearly entities that warrant the theoretical attention of philosophers.

3. Adequacy Criteria

These observations all indicate that we have a widespread practice of discussing and conceiving of ideas as entities in the world; the goal in offering an ontology of ideas is to determine what kind of entities can play these roles. If ideas are to play these roles, including the modest roles attributed to them by ordinary discussions of Einstein's genius or Newton's priority over Leibniz, then any proposed ontology of ideas must meet certain criteria. The criteria proposed here are based on what would be required to play such roles as well as on basic platitudes that seem to govern the way we talk and think about ideas. While we do not take such platitudes to be sacred, and acknowledge that other theoretical virtues—simplicity, elegance, and the like—could trump the fact that a candidate ontology fails to capture them all, it would be a virtue if a theory could preserve as much of our commonsense thought and talk about ideas as possible. Just as an ontology of art should, all else being equal, harmonize with our common practice of engaging with and talking about artworks, an ontology of ideas should accord well with our common practice of engaging with and talking about ideas.

In light of these considerations, we propose four criteria that should be met by any adequate ontology of ideas, the first of which is *Creation*.

> **Creation:** whatever ideas are, they should be entities that we can create (colloquially: *invent*, *come up with*).

Ideas are created by particular people in particular contexts, making them *historical particulars*. (For more on the related and perhaps equivalent notion of *historical individuals*, see Rorbaugh [2003].) Any proposed ontology of ideas that does not satisfy Creation would not be able to make good on the claims that Einstein invented

the idea that gravity is the curvature of space, that Darwin came up with the idea that evolution occurs through natural selection, or that Mozi originally developed an idea of consequentialism in the fifth century BCE, to take just a few examples. Creation, then, is motivated by the desire to straightforwardly understand such claims.

The second criterion is *Distribution*:

> **Distribution:** whatever ideas are, they should be (in principle) *public* rather than *private* entities—that is, entities that can be shared by more than one person.

Many ideas are widely shared, and all ideas can be shared, at least in principle. Any proposed ontology of ideas that does not satisfy Distribution would not be able to make good on the claims that you and I can share an idea of how to preserve Yellowstone Park, that Wilbur and Orville Wright together came up with the ideas central to early powered flight, or that we will someday be able to discover more of the ideas had by various ancient civilizations and so share their insights. As before, Distribution is motivated by the desire to straightforwardly understand such claims.

Third is *Causation*:

> **Causation:** whatever ideas are, they should be entities that are embedded in the causal order.

Ideas can inspire us, motivate us, instill happiness in us, and cause us distress. Ideas, that is, *can make things happen* if only insofar as they are capable of inspiring agents to make things happen. Likewise, ideas can be influenced, shaped, driven, and stunted by other ideas, by the difficult relationship a thinker had with a parent, by class, race, and sex, and so on. Any proposed ontology of ideas that does not satisfy Causation would not be able to make good on the claims that Galileo's ideas about heliocentrism caused a stir in the Church or that the political ideas of Martin Luther King Jr. were substantially influenced by his religious upbringing. If we want to understand such claims straightforwardly, we ought to accept Causation.

Fourth and last is *Evolution*:

> **Evolution:** whatever ideas are, they should be evolving entities in that they are capable of undergoing changes (such as being spread or forgotten).

Unlike, say, numbers, ideas are not static. Instead, like us, they enjoy dynamic life spans. Any proposed ontology of ideas that does not satisfy Evolution would not be able to explain how the idea of free love spread in the United States during the 1960s, agree that David Lewis's ideas about modal realism became more refined between 1971 and 1986, or accept that many of the ideas of the ancient Chinese Yangist philosophers have been lost over time. Ideas can not only change things,

as discussed above—they themselves are also susceptible to change. Adhering to Evolution allows us to make better sense of these claims.

These four criteria will, ideally, be met by any proposed ontology of ideas. Again, it might be that a proposed ontology is attractive enough in other respects that we excuse it for failing to satisfy fully all four of these criteria, but the proponent of such an ontology would then incur the burden of explaining why, given our commonsense thought and talk, ideas seem to have features that their favored view denies them. Wishing to avoid any such burdens, we will take on the task of developing an ontology that fully satisfies all of these criteria.

4. The Systems Account

In aiming to satisfy these criteria, we reject the identification of ideas with *propositions*. When Einstein came up with his idea about gravity, he did not come up with the *proposition* that gravity is not a force, but the curvature of space. That proposition existed long before Einstein existed (in whatever manner propositions exist or exist in time at all). Propositions might be shareable, in that many people can think of the same proposition, but insofar as they are *abstracta*, propositions are causally ineffectual (or at best obliquely causally efficacious) in a way that ideas are not. And then, propositions do not evolve: they are (generally taken to be) eternal and unchanging. When there is a change in how many people believe a proposition, for instance, that is normally taken to be a change in the people, not a change in the proposition. So whatever we take ideas to be, we should not take them to be propositions. (Similar arguments can be made that ideas should not be identified with senses, facts, conceivable states of affairs, or other such entities. For further arguments as to why ideas should not be identified with propositions, at least in the context of the philosophy of late modern art, see Schellekens 2005: 20–21.)

Is the above argument that ideas should not be identified with propositions too quick? Consider the claim that Leibniz and Newton came up with distinct ideas. One might well deny this. After all, they both came up with an idea *of* the *same thing*: the calculus. Is it really so wrong to hold that ideas are *what* we think about and hence are really entities that serve as *contents*, such as propositions?

We grant that it is quite natural to talk as if Leibniz and Newton independently came up with the same idea and in general to talk about ideas as the propositions toward which we have attitudes. But it is also natural to talk about ideas as the attitudes toward these propositions. Consider: if we each buy our own copies of *Finnegans Wake*, there is certainly a strong sense in which we have purchased the same book: we purchased tokens of the same book type. There is an equally strong sense in which we did not: we purchased two distinct tokens. Along the same lines, Leibniz's idea and Newton's idea are both tokens of *some* type: *idea associated with such-and-such propositions of mathematics*. So, while there certainly is a very strong sense in which Leibniz and Newton had the same idea—their ideas were tokens of the same idea type—there is also an equally strong sense in which they did not—they created distinct token ideas. The intuition that Leibniz and Newton independently came up with the same idea is accounted for by focusing

our attention on the former sense of 'same' (the type sense), rather than the latter (the token sense).

So far, this suggests only that there are two ways of thinking about ideas that are, at best, on a par. So why proceed as we do? Individuating ideas in our manner is necessary for explaining one of the phenomena that got us started with the investigation of ideas in the first place: the thought that Newton kept his mathematical ideas to himself, forcing Leibniz to invent his own and leading to a quarrel over whether Leibniz really invented his idea or merely stole it somehow from Newton. This way of thinking about things is made possible by focusing our attention on the token sense of 'same idea'. Likewise, the genius of Einstein's idea is not in its content but in Einstein's grasping of that content in the intellectual context he did. When a student today grasps that content while taking a physics course, the student's grasping of that content does not display the same genius. The history of ideas, historiography, the philosophy of art, and the rest are all likewise interested, not in the eternal propositions that have been grasped by people, but in the grasping of these propositions, in who first grasped them, and in the distribution, evolution, and causal roles of these graspings. So, if Leibniz and Newton engaged in distinct, causally isolated graspings, then we think it best to say that Leibniz and Newton generated different ideas. If, on the other hand, it turns out that one was sufficiently causally and historically influenced by the other's idea, then we have one idea, the credit for the origination of which is in dispute.

We conclude, then, that ideas should not be identified with propositions, contents, or the like, not because doing so violates all the ordinary ways we have of talking about ideas, but because doing so makes it difficult or impossible to ask or answer the sorts of questions we began with—questions that are about ideas.

A better approach, and our proposal, is to hold that ideas are *systems of causal-historically related token mental states of the same type*. An idea is generated when a person (or group of people) comes to have a novel token (contentful) mental state. Einstein generated his idea about gravity when he first had a thought (a token mental state) with the content that gravity is not a force, but the curvature of space. (We assume that whatever the contents of mental states might be—propositions, senses, states of affairs, or what have you—they are not tied to language. Einstein expressed his thought's content in German, but the same thought content can be expressed in English as well.) At the moment of an idea's generation by one person, it is made up of the single token mental state in that person's head. There are some interesting complications regarding ideas that are so complex that they cannot be grasped all at once and require external aids (such as a score, for grasping how a whole symphonic orchestra might proceed to play in a novel way). We will set them aside for now. If we can adequately account for simpler ideas in the present paper, that will be progress enough.

Einstein's idea spread through the physics community as it was communicated to others who then themselves came to have tokens of that same mental state type (that is, token mental states that share content). Imagine that, for the first week after inventing his idea, Einstein communicated it to no one except Mileva Marić. At that time, the idea was the *system* made up of Einstein's token mental state and Marić's token mental state. (We leave the notion of *systems* loose and intuitive for

now, tightening it up in the next section.) As more people came to have tokens of that mental state type, derived ultimately from Einstein, the system grew. This is what it is for an idea to *spread*. As people lose the relevant mental states, the system shrinks, and the idea comes to be *forgotten* in certain places or in certain communities. When the system shrinks until the very last mental state in the system is gone and none can ever come back, the idea is entirely lost.

We write 'mental state' because we want to remain mostly neutral about just what mental states can form systems that count as ideas; it seems quite a few can. A person with an idea about the shape of space has a thought or a belief, but a person with an idea for how to invade Troy has a plan, which, as Bratman (1987) argues, might be distinct from thoughts and beliefs. A person with a desire for a certain sort of professional career would seem to have an idea of what she wants, as well. Nonetheless, we take it that the ideas we are trying to explain are concrete attitudes with complete propositions (or some similar entities) as contents. (Philosophers with other ideas about the semantic features of particular mental states should feel free to deploy them here in place of our talk about propositions.) In the jargon of some philosophers, having the concept of a cat is having an idea, but for our purposes just knowing that you have the concept CAT does not yet tell us what ideas, if any, you have about cats.[1]

We assume that mental states are physical states. Ideas conceived of as systems of token mental states—which themselves would be something like particular arrangements of brain matter—are thus concrete particulars with the potential to be spatially discontinuous. We describe ideas as *potentially* spatially discontinuous, rather than just spatially discontinuous *tout court*, since a system with only one token will be spatially continuous until more tokens are added to the system.

Given that they are physical entities, Einstein and Marić's token mental states presumably exist in proper subregions of the spatiotemporal regions occupied by Einstein and Marić, respectively. The idea itself is located where the token mental states are located—perhaps wholly, in the same way that universals are said to be wholly located wherever they are instantiated, or perhaps partially, in the same way that I am partially located wherever one of my parts is located. While we here take no official stance on this matter either way, and the account on offer is ultimately compatible with either position, we admit a tentative preference for the latter approach if only because partial location is easy enough to grasp, whereas the burden of making good sense of an entity's being wholly located in more than one location is perhaps an onerous one.

In taking ideas to be (potentially spatially discontinuous) concrete particulars, we make good sense of other ways of talking about ideas. Talk of an idea *spreading* or *being lost* often takes on explicitly spatial connotations: the missionaries spread their ideas to the new land; the idea of hook-and-line fishing disappeared from Tasmania after rising seas cut its land bridge to the rest of Australia; it took a long time for the new political ideas to reach New France, and so on. Such talk, on our account, can be taken literally.

---

[1] Thanks to Kevin Scharp for pressing us to be more clear here.

Of course, some will reject the claim that mental states are best thought of as physical states. Thankfully, if mental states are nonphysical, this would require only slight changes to what we've said so far. If mental states are dispositions or normative statuses, for instance, these dispositions or normative statuses will still be grounded, in part, in token physical states found inside the skulls of the people with the mental states, and that will suffice to make sense of what we say here. Similarly, if Cartesian dualism holds, then the causal features of ideas will be mediated by their relations to neural states, and it will be these neural states, strictly speaking, that have locations rather than the ideas themselves.

If ideas are systems of mental states, then what happens when, say, Einstein wrote his idea down? It doesn't seem that he wrote down a system of mental states, physical or otherwise. Rather, he wrote down a sentence with the same content as the content of one or more of his mental states. However, people often talk about ideas by talking about the contents of the beliefs (or other attitudes) that make up the idea. The idea of a perpetual motion machine is said to be impossible, but it is not the idea that is impossible: it is the associated content that would be impossible to realize. The idea that a small team should cross the enemy lines and sabotage the incoming supply train is said to be a daring idea, but the idea itself is daring only if the person proposing it expects to be part of the small team; otherwise, the idea calls for daring behavior rather than exemplifying it. We say that Einstein wrote his idea down, then, when he inscribes a sentence that shares content with his mental state that is the idea or is part of it.

We call this account the *Systems Account*: an idea is a *system* of token mental states that (1) are of the same type in that they share the same content, and (2) are connected through appropriate causal-historical relations. Insofar as token mental states are physical entities, systems of such entities will also be physical entities; ideas, then, are physical entities. The content of a particular idea is the content had by the token mental states in the system.

In allowing ideas to have origins in space and time, the Systems Account satisfies Creation: an idea is created when and where the first mental state in the system is tokened. Thus, if Leibniz and Newton's ideas originated independently in a causal-historical sense, then those ideas were created by different people in different contexts and, hence, distinct. Furthermore, the creativity, naivety, class struggle, or the like expressed in the creation of the first token mental state in the system are thus the creativity, naivety, or class struggle expressed in the invention of the idea. In cases in which a group of people collectively generates the first relevant mental state token, the people involved can be said to have collectively created the idea.

In taking ideas to be concrete particulars, albeit potentially spatially discontinuous ones, the Systems Account satisfies Distribution. Ideas can be shared (by tokening the requisite mental state type, derived from an existing system), investigated by people not sharing them, and otherwise treated as shared public (while also mental) entities.

The Systems Account also satisfies Causation: making sense of how ideas can inspire us, make us happy, frighten us, and so on, is now no more complicated than making sense of how mental states and systems of shared mental states can cause

inspiration, emotions, motivations, and the like. Similarly, ideas can be inspired by conversations, serve as expressions of religious feeling, or be a product of their times. Some philosophers of mind take it to be quite a puzzle how mental states could be involved in these causal relations, of course (e.g., Kim 1998). But this is no more a puzzle about ideas than about, say, desires for food, and we take this to be enough of a reduction in puzzlement for now.

Finally, the Systems Account satisfies Evolution: ideas are spread as more of the right kind of token mental states enter into the system, they diminish as fewer token mental states of the same type participate in the system, and are forgotten when the system loses its last token mental state. As for evolution of contents, here is one way to understand this within the Systems Account: sometimes the content of a mental state is quite complex. The plan for making a big holiday meal might have a very large number of steps, for instance. Such a plan can evolve in that its content is almost the same before and after the putative evolution (the same meal will appear on the table, for instance, and the sequence of many of the steps remains the same) but differs in select particulars (the sprouts will be purchased separately, from a farmers' market, not from the large grocery store intended originally). This can count as the same idea if one types *sameness* by something looser than perfect identity of content, but since this is just how people talk about certain ideas it seems reasonable: people say 'I tweaked the idea' rather than 'I abandoned the idea and came up with a very similar but strictly distinct new one'. In these contexts, some combination of a causal-historical relation to the original idea, plus substantial overlap in content, is treated as sufficient for calling the new plan (theory, etc.) a new version of 'the same idea'. Perhaps this case does not cover all possible modes of content evolution, but it at least suggests that such evolution makes sense in principle on the Systems Account.

5. An Ontology of Systems

The metaphysical plausibility of the Systems Account depends in large part on what we take *systems* to be. In this section, we consider various strategies for characterizing the nature of systems. Our goal in this section is not merely exploratory, but is instead to weigh the various advantages and disadvantages, if we take *ideas* in particular to be systems of certain mental states, in identifying these systems with various, more familiar entities from the metaphysician's toolbox.

5.1 Systems as Sets

The question that faces us is the question of what *systems* are. To start with, we might take systems to be *sets* of mental states, unified into a set in virtue of both the type of content they have and their causal-historical relations. But this will not do. A system of mental states, whatever it is, is a dynamic, concrete, causally efficacious, historical individual—it is located in space and time, created in a particular context and developed from there via a specific causal-historical chain

of events. The set, by contrast, is static and abstract, not located in space, and not created in a particular context. (For an account according to which abstract types *can be* created, see Levinson 1980; for further relevant discussion see also Caplan and Matheson 2004.) The *members* of the set have spatiotemporal properties and so on, but the set embracing them does not, and so should not be identified with the system of mental states making up an idea. Similar concerns militate against identifying systems with any other kind of abstract object (classes, types, etc.). Whatever ideas are, they cannot be sets (or classes, types, etc.) of token mental states.

5.2 Systems as Systems

A more metaphysically profligate approach to our account of ideas is to take the Systems Account at its word and to ground a metaphysics of ideas in a metaphysics of *systems*—without further giving an account of systems in terms of any other, more familiar metaphysical posits. On this view, systems are metaphysical posits in their own right:

> **Systems:** For any concrete objects of type x and any relation between them R, there is a *system* of x's related by R. This system is an entity that (1) comes into existence when the first of the x's comes into existence and goes out of existence when the last of the x's goes out of existence, and (2) has all the x's related (in some way) by R.

Note that, on this formulation, a system of some (appropriately related) x's might exist at t even if none of those x's exists at t, as long as at some later time $t*$ there will again exist an (appropriately related) x. This consequence will be discussed further later on, in section 6.

Understood in this way, many things would count as systems. The first thirty ten-dollar bills printed by the US Mint after midnight on January 1, 2016, related to one another by temporal precedence, form a system. The male-line descendants of Genghis Khan such that each one is loved by at least one other such descendant form another system. And so do the animals related to one another by descent from one originating pair: the lineage of *Homo sapiens* starting from some African Adam and Eve, perhaps.

There are many systems of mental states conceived in this way, and an idea will be a system of a specific sort. It will have some mental states, each with the content P (or closely related family of contents P_1, P_2...), related to one another by the relation of *having-been-caused-by-grasping-the-existence-and-content-of*, back to the originating mental state that will have been caused in some other way and so be (in some measure) original. The system of the x's is located (again: perhaps partially, perhaps wholly) where each of the x's is located, and so ideas are located wherever people with the right mental states (plans, desires, etc.), derived from the right originating thinker (planner, wisher, etc.), are located. Understanding ideas as systems of mental states in this sense avoids the problems that come with

understanding ideas as sets of such things, making this a more promising approach than the previous one.

5.3 Systems as Fusions

Though this *sui generis* approach to systems and so to ideas is straightforward, it would not be unreasonable to complain that it just adds needlessly to our pile of ontological posits. A less metaphysically profligate strategy might be to identify systems of mental states—and so ideas—with certain *mereological fusions*. Doing so would allow ideas to be concrete, historical individuals while also being things that have a reasonably well understood metaphysical foundation.

Based on work by Kit Fine (1994), we can consider two different kinds of fusions: *compounds* and *aggregates*. One way of characterizing the difference between compounds and aggregates is by looking at their persistence conditions: a compound exists at *t* only if *all* of its parts exist at *t*, whereas an aggregate exists at *t* only if *at least one* of its parts exists at *t* (cf. Caplan, Tillman, and Reeder 2010: 503). So, if ideas are fusions, are they compounds of token mental states or aggregates of such things?

We favor ideas as aggregates. Einstein's idea begins with Einstein, but it has outlived its creator and will outlive the authors of this essay. The token mental states in virtue of which the idea existed in 1905 are just a small fraction of the totality of token mental states making up the system of token mental states that is the idea—yet the idea existed then, which could not be true if ideas were compounds. Similarly, the idea exists now even though only a fraction of the token mental states making up the totality of the system exist now: Einstein's token mental state, sadly, no longer exists, and there are many future token mental states that have not yet come to exist but that soon enough will and then become part of the system. Aggregates of token mental states, where the token mental states that compose the aggregate are selected on the basis of their descent (by intellectual grasping) from some original token mental state that is not so descended from any other token mental state of the same type, are well defined entities that seem to have the persistence conditions required.

Aggregates are historical particulars, coming into existence at one time, entering into causal relations, changing, and eventually passing out of existence. As such, they are well suited to satisfy Creation, Causation, and Evolution. And given that the aggregate is a fusion of mental states found in different people's heads, Distribution seems readily satisfied by them as well. Thus, aggregates would appear to offer a convenient metaphysical reduction of ideas from irreducible systems to fusions of individual mental states tied together by their causal and historical relations to one another.

Before moving on, a few notes on this approach are in order. First, depending on your view of when composition occurs, there might be *many* aggregates of mental states that do not qualify as ideas. As presently conceived, ideas are *certain* aggregates of *certain* mental states—that is, they are aggregates of mental states that are related, causally and historically as well as by content, in the right way.

Second: if parthood is transitive, the friend of the fusions strategy is committed to saying that since ideas are systems of mental states and systems are fusions, then if mental states have proper parts (perhaps, bits of brains), then those proper parts are also parts of the idea. While we're not sure that we have strong intuitions either way about what does and does not count as a part of an idea, we take it to be a potentially viable move for the friend of the fusions strategy simply to deny the transitivity of parthood. Those committed to the transitivity of parthood might be better served by the approach discussed previously in section 5.2 and the approach to be discussed in section 5.4.

Finally, we have relied on Fine's notion of *aggregate*. Though this notion can be distinguished from an earlier notion under the same name introduced by Tyler Burge, present purposes do not require that we distinguish between them. On Burge's (1977: 102) account,

> [a]ggregates, like their member-components and unlike sets, are physical objects having spatio-temporal location and the capability of exerting forces and undergoing changes. Like their member-components, aggregates come into and pass out of existence. Member-components of aggregates are spatio-temporal parts of the aggregate, although not all parts of an aggregate are member-components of it. For example, each tiger in the aggregate of all tigers is part of that aggregate; but whereas proper parts of tigers (tiger hearts) are parts of that aggregate, they are not member-components of it.

As per this characterization, our four adequacy criteria are satisfied whether we take ideas to be systems of mental states as Burge-aggregates or as Fine-aggregates. As such, throughout the remainder of the discussion, all discussion of aggregates is intended as neutral between the two characterizations.

In sum, while taking ideas to be systems *qua* compounds is not feasible, there is some promise in taking them to be systems *qua* aggregates, in either a Finean or Burgean sense.

5.4 Systems as Pluralities

Another even more ontologically conservative option is to take systems to be *pluralities*: referentially singular collections of objects, such as those discussed in Boolos (1984). There are some papers on the desk: there is a set of those papers, and there is (perhaps) a fusion of those papers, but we also have *the papers*. Likewise, whenever there are some mental states, there is also a set of those mental states and (perhaps) a fusion of those mental states, but there are also *the mental states* themselves. If ideas are systems of (appropriately related) mental states and systems are best thought of as pluralities, then an idea would be *some (appropriately related) mental states*.

Taking systems to be pluralities rather than aggregates comes with an advantage: talk of pluralities is most often taken to be ontologically innocent, whereas talk

of aggregates is not. A person who accepts an aggregate of some x's takes on a commitment to an object—namely, the aggregate—composed of those x's; in contrast, a person who accepts *the x's* as a plurality incurs no ontological commitment beyond commitment to the x's. The pluralities strategy thereby enjoys, to a greater extent than the aggregates strategy, the theoretical virtue of *parsimony*. More specifically, it is more *qualitatively ontologically parsimonious*: it comes with minimal ontological commitments to *kinds* of entity. (For more on qualitative ontological parsimony, see Lewis [1973: 87 and 1986: 2].) A commitment to *systems* of entities, so conceived, is no more than a commitment to pluralities of those entities, which is, in turn, no more than a commitment to those entities (and certain relations between them). In the case of ideas, rather than positing ideas as entities in addition to mental states, the pluralities strategy incurs no ontological commitment beyond mental states (and causal-historical relations, contents, etc.). For those of us who already carry an ontological commitment to such entities, ideas *qua* pluralities come for free. We take this to be a serious advantage of this strategy.

Despite this advantage, there are problems with the pluralities strategy as well. Begin with Creation. Einstein's idea about the nature of gravity was created in 1905 by Einstein. But the mental states making up the system that is Einstein's idea were not all created in 1905. Since Einstein's idea has been widely shared and continues to be shared with new people every day, some of the mental states in the system did not exist until earlier this week. It seems hard to say that *the mental states* were created in 1905; the first mental state in the system was created then, but not all the mental states were created then.

A defender of the pluralities strategy might respond that all the relevant mental states came, in the causal-historical sense, from the first mental state back in 1905, but this might not be enough. One wants to say not just this, but also that the *idea* was created in 1905. This seems to require something less than saying that the *mental states* were all created in 1905, suggesting that the idea is something different from the (relevant plurality of) mental states.

Causation also appears to be a problem. We hesitate to generalize from '*that x* caused this-or-that' to 'the x's caused this-or-that'. Suppose a bunch of rocks fall from the sky, and one of them knocks you out. The rock that knocked you out is one of *the rocks*, but *the rocks* didn't knock you out—*that* rock did. The advantage here would go to the aggregate strategy: the aggregate of the rocks knocked you out, in virtue of one of its parts knocking you out. So, if we have to choose between the aggregates strategy and the pluralities strategy and we want to hold on to causation, we should go with the aggregates strategy.[2]

The defender of the pluralities strategy might respond, however, by questioning the supposed data. Suppose that a group of philosophers are in hotel bar, but throughout the whole night only one of them purchases a drink. The bartender, frustrated, might say '*the philosophers* ordered only one drink!'[3] This seems a correct thing to say, and the corresponding causal claim, '*the philosophers* caused

---

2 Thanks to Dan Korman for raising this challenge.

3 We owe this example to Ben Caplan.

only one drink to be poured', also sounds fine to us. But that is an instance of generalizing from 'that x caused such-and-such' to 'the x's caused such-and-such'. If this is right, then the pluralities strategy might be in the clear with respect to satisfying Causation.

When it comes to what is *possible* for ideas, though, things again look bad for the friend of the pluralities strategy. Suppose that two people share a plan for grooming a dog. It seems that the idea could have been shared by a third person. But is it equally clear that those two mental states could have been these three mental states? That seems odd.

What seems odd in some contexts seems less odd in others, however. Could these dozen people making up the wedding party have been a simpler trio? We think this makes some sense, and yet it seems very strange indeed to say just that these twelve people could have been those three; what seems more natural to say is that *the wedding party* could have been made up of three people though in fact it comprised a dozen. And perhaps what this suggests is just that wedding parties—and ideas, and many other kinds of entity—are most naturally treated as amounting to something more than just the plurality of individuals actually composing them.

Thus, we tentatively suggest that ideas are not best understood as systems *qua* mere pluralities of (appropriately related) mental states. For now, it seems better to treat ideas as systems *qua* irreducible systems (as in section 5.2) or as systems *qua* aggregates (as in section 5.3).

6. Temporal Discontinuities

We finish our main discussion with a metaphysical puzzle that will arise for any account of ideas: what is one to say about the seeming ability of ideas to survive gaps in mental states? Suppose Octavio comes up with the idea of calling Norma 'cuddle bunny'. His idea is not particularly original, but it might be the first time anyone thought of using this term of endearment for Norma (and even if not, Octavio might have invented it independently). And suppose Octavio writes a note on the back of a photo using this term, making it clear to the reader that he had this idea. If Octavio and Norma part and go on to have separate lives, it might eventually happen that they and everyone else forget Octavio's idea. But then, perhaps generations later, someone might find the photo and read the note and recover Octavio's idea. In this situation, the idea of calling Norma 'cuddle bunny' has not been reinvented independently; rather, *Octavio's idea* has spread to a new person—the archivist, perhaps, handling the photo. How can this apparent power of ideas to survive periods of nonexistence be explained?

In answering this, we find a brief detour through the metaphysics of species illuminating. It is not uncommon to take species to be individuals with particular organisms as parts of the larger whole (for example, in Hull 1978). It seems problematic, however, to say that species are compounds of particular organisms. After all, intuitively, a species can exist at t even if not all of the relevant organisms exist at t. (If one is a *presentist* in the sense that one holds an ontological commitment only to presently existing entities, then the only organisms that *could*

be part of a species-compound at *t* are organisms that exist at *t*.) Compounds, however, are fragile: a change in part necessitates a change in compound. Taking species to be compounds of organisms, then, comes with the consequence that when an organism dies, its species goes with it. We might conclude, then, that species must instead be aggregates of particular organisms.

But this might be problematic too. Again, an aggregate exists at *t* only if at least one of its parts exists at *t*. But suppose that all of the stegosauruses are dead. Suppose further that we have, well-preserved in a mosquito trapped in amber, some *Stegosaurus armatus* blood containing fully intact DNA and that it is within the realm of technological possibility to use that DNA to clone a new stegosaurus. It might not be implausible to claim that in such a scenario the species of *Stegosaurus armatus* is not extinct after all. (If this scenario is too fanciful, consider that the role of seed banks is to allow species of wheat, corn, and the like to persist in some manner even if disease kills every living plant. Structurally, the reintroduction of wheat from seeds would be the same situation.)

Certainly, if scientists go ahead and make more members of a species, it will be shown that the species is not extinct since there will be extant members of the species! What should then be said about the existence, not of individual animals, but of the *species* in the time between the death of the last member of *Stegosaurus armatus* long ago and the birth of the first new one? And what should be said about the species if the power to make a new member of *Stegosaurus armatus* exists but is never used?

Three approaches suggest themselves. The first approach holds that the species can survive times during which it does not exist. The species exists, goes out of existence temporarily, and then comes back into existence. The second approach holds that so long as there is (or will be) some future member of *Stegosaurus armatus* that exists, then the species continues to exist through any gaps in the existence of individual members. And the third approach holds that so long as a future member of the species *can* come into existence, the species continues to exist, regardless of whether any of the animals now or ever will live.

Taking either the second or third approach requires giving up on the idea that the species *Stegosaurus armatus* is an aggregate of particular organisms since that aggregate does not exist at times when none of its parts exists. But these two approaches make it easy to talk about extinction: a species has gone extinct either when the last organism in the species has died (on the second approach) or when there can be no more organisms in that lineage (on the third approach).

Taking the first approach, on the other hand, allows one to continue to hold that a species is simply an aggregate of individual animals (related by common descent) but makes it harder to talk about extinction. A species can seemingly go extinct, then cease to be extinct, and then go extinct again. And this seems rather odd. Perhaps the best thing to say, on the first approach, is that extinction is something stronger than mere (contingent, perhaps temporary) nonexistence for a species. Rather, extinction is *terminal* nonexistence; temporary nonexistence of a species would then need some other name, such as 'quiescence'. On this way of thinking, then, *Stegosaurus armatus* might be quiescent if we will eventually make more members of the species, but it might be extinct if we never will do so.

Let us return now to ideas. It seems that the same three approaches can be taken and that the same patterns of advantage and disadvantage apply. The first approach would be to hold that ideas can come into and go out of existence many times once they have been invented; this allows one to continue thinking of ideas as aggregates of mental states. The second approach would be to hold that ideas can only cease to exist once: they exist even when no one grasps their contents (in the appropriate causal-historical chain) so long as someone will eventually grasp the relevant content in the appropriate way once again. And on the third approach, ideas only cease to exist when it becomes impossible for someone eventually to share in the idea again; the causal-historical chain has necessarily ended. On the second and third approaches, though, it would not be possible to think of idea systems *qua* aggregates of mental states since ideas would exist even when no part of the aggregate exists.

If the metaphysical parsimony of reducing systems of mental states to aggregates of mental states seems important to the reader, then only the first approach should seem acceptable. Is this a problem? It might seem so since it requires saying that for years the idea of calling Norma 'cuddle bunny' did not exist, and then it came to exist once again when the archivist found the old photo. And that strikes us, at least, as a puzzling thing to say.

To ameliorate this puzzlement, though, the philosopher favoring the first approach might point out that it is common to say that some ideas are *merely forgotten* whereas others are *permanently lost*. The idea of calling Norma 'cuddle bunny' is merely forgotten until the archivist finds the photo; if the photo (and other traces) had all been destroyed, though, the idea would be genuinely lost forever. The jargon we might use for the underlying facts, the advocate of the first approach might say, could be as follows. Systems of mental states that have current members are systems such that *the idea exists now*; systems that have no current members but that have future members are *temporarily forgotten*, that is, they are not currently existing but will exist again in the future. Systems that have no present or future members are systems such that the idea is *permanently forgotten*, that is, it will never again exist. And systems that can have no present or future members are systems such that the idea is not just forgotten, but *lost*, that is, it can never again exist.

An idea that has the property of *being forgotten* might nonetheless appear to be an idea that must exist, or else it would not be in a position to have that property. And if this seems reasonable, then the second or third approach will be preferred.

We face a fork in the road. If, on the one hand, we are comfortable with ideas being potentially temporally discontinuous, coming in and out of existence as people enter the relevant mental states, then the aggregates strategy should be sufficient. If, on the other hand, we continue to hold that ideas *do* stay in existence even when there are no relevant mental states, we will need to identify systems with something else: with *sui generis* systems. We leave this point open; the best way to proceed appears to depend on how important it seems to be to reduce systems to aggregates instead of identifying them with systems *qua* systems as discussed in section 5.2, and this is an issue that strikes us as substantially deeper than how to think specifically about ideas.

7. Further Work

So far, we have motivated the desire for an ontology of ideas and worked toward specifying what conditions an adequate account would have to satisfy. We then offered our own account, the Systems Account, according to which ideas are (potentially spatially discontinuous) systems of causally and historically related token mental states of the same type. Depending on our ontological scruples and whether we are willing to allow ideas to be not just potentially spatially but also potentially temporally discontinuous, we might take this talk of *systems* to amount to talk of aggregates, pluralities, or a new kind of entity altogether.

There is, of course, much work left to be done. We have repeatedly invoked the notion of *appropriate causal and historical relatedness*, but never specified exactly what this relatedness might amount to. While we would like our account not to turn on some particular understanding of this notion and instead be compatible with many potential realizations of appropriate causal and historical relatedness, we think that questions about such relatedness are both interesting and important, especially if ideas are to play the various roles discussed in section 2.

Earlier, we mentioned briefly the intuition that as dynamic entities ideas can *evolve* over time. What conditions must a token mental state satisfy to qualify as not a new member of an extant system and, hence, an expansion of an extant idea, but rather as the first member of a new system and, hence, the beginning of a new idea? Some ideas seem to change over time: is this phenomenon best explained by pointing to a change in the content of the relevant mental states in the system or perhaps by taking distinct ideas at different times to be similar enough and appropriately causally and historically related, such that we apply some loose notion of 'sameness' to them? These are interesting and important questions that are also relevant to the roles discussed in section 2.

Finally, one might wonder what to say in cases in which distinct ideas seem to have merged over time. Suppose it is true that Leibniz and Newton both came up with their respective ideas of calculus in the relevant kind of causal-historical isolation from one another. Since then, students have come to be taught what seems to be *one* idea of calculus, rather than what is, on our account, *two*. Perhaps there is a way to make sense of this case as one in which the two ideas have somehow converged into one idea over time, or perhaps there is now a third idea of the calculus being taught, influenced by but strictly distinct from the original two ideas of Leibniz and Newton. Alternatively, we might say that students are learning just Leibniz's idea or just Newton's idea or some bits of both, or we might simply concede that students are learning two ideas but merely recognizing them as one. This is a difficult issue, and the answer will turn on how we might choose to address some of the issues left unresolved in this paper.[4]

To echo what was stated initially, however, our goal in this discussion has been modest: we have aimed to introduce and motivate discussion on the ontology of

4 Thanks to an anonymous referee for raising this issue.

ideas and offer a promising account of our own. There is further work to be done, but that further work is exactly that.

WESLEY D. CRAY
GRAND VALLEY STATE UNIVERSITY
crayw@gvsu.edu

TIMOTHY SCHROEDER
RICE UNIVERSITY
timothy.allan.schroeder@gmail.com

References

Alberro, Alexander, and Blake Stimson, eds. (1999) *Conceptual Art: A Critical Anthology*. Cambridge, MA: MIT Press.

Boolos, George. (1984) 'To Be is to Be the Value of a Variable (or to be Some Values of Some Variables)'. *Journal of Philosophy*, 81, 430–49.

Bratman, Michael. (1987) *Intentions, Plans, and Practical Reason*. New York: Cambridge University Press.

Burge, Tyler. (1977) 'A Theory of Aggregates'. *Nous*, 11, 97–117.

Caplan, Ben, Chris Tillman, and Patrick Reeder. (2010) 'Parts of Singletons'. *Journal of Philosophy*, 107, 501–33.

Caplan, Ben, and Carl Matheson. (2004) 'Can a Musical Work Be Created?'. *British Journal of Aesthetics*, 44, 113–34.

Collingwood, R. G. (1938) *Principles of Art*. London: Oxford University Press.

Cox, Renée. (1986) 'A Defence of Musical Idealism'. *British Journal of Aesthetics*, 26, 133–42.

Cray, Wesley. (2014) 'Conceptual Art, Ideas, and Ontology'. *Journal of Aesthetics and Art Criticism*, 72, 235–45.

Croce, Benedetto. (1909) *Aesthetics*. Translated by D. Ainslie. London: Vision Press/Peter Owen.

Davies, David. (2008) 'Collingwood's "Performance" Theory of Art'. *Journal of Aesthetics and Art Criticism*, 48, 162–74.

Dodd, Julian. (2007) *Works of Music: An Essay in Ontology*. Oxford: Oxford University Press.

Everett, Anthony, and Timothy Schroeder. (2015) 'Ideas for Stories'. In Stuart Brock and Anthony Everett (eds.), *Fictional Objects* (New York: Oxford University Press), 275–93.

Fine, Kit. (1994) 'Compounds and Aggregates'. *Nous*, 28, 137–58.

Goldie, Peter, and Elisabeth Schellekens, eds. (2010) *Who's Afraid of Conceptual Art*. New York: Routledge.

Goldie, Peter, and Elisabeth Schellekens, eds. (2007) *Philosophy and Conceptual Art*. Oxford: Clarendon.

Hight, Marc. (2008) *Ideas and Ontology: An Essay in Early Modern Metaphysics of Ideas*. University Park, PA: Pennsylvania State University Press.

Hull, David. (1978) 'A Matter of Individuality'. *Philosophy of Science*, 45, 335–60.

Kim, Jaegwon. (1998) *Mind in a Physical World*. Cambridge, MA: MIT Press.

Kosuth, Joseph. (1969) 'Art After Philosophy'. *Studio International*, 178, 134, 160–61, 212–13. Reprinted in Alberro and Stimson (1999): 158–77.

Levinson, Jerrold. (1980) 'What a Musical Works Is'. *Journal of Philosophy*, 77, 5–28.

Lewis, David (1986) *On the Plurality of Worlds*. Malden, MA: Blackwell.

Lewis, David. (1973) *Counterfactuals*. Malden, MA: Blackwell.

LeWitt, Sol. (1969) 'Sentences on Conceptual Art'. *0–9*, 5, 3–5. Reprinted in Alberro and Stimson (1999): 106–08.

Ridley, Aaron. (1997) 'Not Ideal: Collingwood's Expression Theory'. *Journal of Aesthetics and Art Criticism*, 55, 263–72.

Rorbaugh, Guy. (2003) 'Artworks as Historical Individuals'. *European Journal of Philosophy*, 11, 177–205.

Schellekens, Elisabeth. (2005). '"Seeing is Believing" and "Believing is Seeing"'. *Acta Analytica*, 10, 10–23.

Trerise, Jonathan. (2010) 'Against the Strength of Patent Protection'. *The Monist*, 9, 464–80.

Wollheim, Richard. (1980) *Art and Its Objects*. 2d ed. New York: Cambridge University Press.