

(2018, DOI: 10.1007/s10670-018-0097-2)

“Twofileness. A Functionalist Approach to Fictional Characters and Mental Files”,

Erkenntnis

Twofileness. A Functionalist Approach to Fictional Characters and Mental Files

This paper considers two issues raised by the claim that fictional characters are abstract artifacts. In making this claim, Amie Thomasson (1999) has characterized abstract artifacts as products of human creativity that, unlike concrete artifacts, lack a spatial location but have a beginning and possibly an end in time. If one endorses this view there are two basic questions that should be addressed. First, given that artifacts normally have functions, what is the function of a fictional character? Second, given that, in experiencing works of fictions, we usually treat fictional characters as concrete individuals, how can such a phenomenology fit with an ontology according to which fictional characters are abstract artifacts?

I will indirectly address the second issue by directly addressing the first one.

In §§ 1-2, I will claim that the function of fictional characters is the generation of mental files of a special kind. I will argue that our experience of fictional characters as concrete individuals depends on the “irregular” mental files that are generated by fictional characters as abstract artifacts. Furthermore, I will show that appreciators of works of fiction can open not only “irregular” files about fictional individuals but also “regular” files about the corresponding abstract artifacts. This is the phenomenon that I will call ‘twofileness’.

In § 3, I will address the objection according to which mental files about fictional individuals are not generated by fictional characters themselves, understood as abstract artifacts, but simply by works of fiction. I will show that one can conceive of fictional characters as the functional components of a work of fiction that generate such mental files. From this perspective, fictional characters contribute to the functioning of a work of fiction much as the respiratory system or the digestive system contribute to the life of an animal. I will argue that this is a good reason to endorse artifactualism about fictional characters. I will call this ‘the argument from mental files’.

In § 4 I will show that this argument is compatible with both an “austere” artifactualism according to which fictional characters are nothing but functional components of works of fiction

and a “robust” artifactualism according to which fictional characters are self-standing entities that are embodied in works of fiction. Although the argument from mental files helps us to highlight the distinction between these two kinds of artifactualism, it does not allow us to choose between them. It remains neutral on this point.

In § 5, I will clarify the twofileness phenomenon by figuring out the relationship between the two kinds of mental files that are at play in our engagement with fiction, namely mental files about concrete individuals in the fictional world and mental files about abstract artifacts in the actual world. In so doing, I will show that twofileness can play a role not only in our engagement with fiction but also in our thinking or speaking of fictional characters from the outside of such an engagement.

In § 6, I will address another objection to the argument from mental files. According to this argument, one should conclude that historical figures, in turn, exist not only as concrete individuals but also as abstract artifacts. This is clearly absurd and therefore – so the objection runs – the argument should be wrong. I will reply by biting the bullet, that is, by showing that there is no absurdity in treating historical characters as those functional components of works of history that allow us to deploy mental files about the corresponding historical individuals. Yet, there is a crucial difference between fiction and non-fiction in this respect; in the case of historical individuals, one mental file is sufficient in order to store all the information we can acquire about them whereas, in the case of fictional individuals, two files are needed. In this sense, the notion of twofileness can help us to shed some light on the distinction between fiction and non-fiction.

In § 7, I will consider one last objection to my functionalist approach according to which fictional characters are abstract artifacts whose function consists in generating mental files in the minds of appreciators of works of fiction. According to this objection, fictional characters, as *abstract* artifacts, cannot originate causal chains that end up in the generation of mental files in the minds of appreciators. I will reply to this objection by focusing on the process of creation of

fictional characters. I will argue that the abstract objects created by such a process can originate causal chains that pass through concrete tokens and end up in the minds of appreciators.

Finally, in § 8, I will draw my conclusions.

1. Drawing on the seminal works of Paul Grice (1969) and Peter Strawson (1974), François Recanati (2012) defines a mental file as a vehicle of singular thought whose function is to store *information* acquired by a subject in virtue of some “epistemically rewarding” *relation* to a particular individual. In short, a mental file is a pair $\langle \textit{information about the individual}, \textit{relation to the individual} \rangle$. The relation is “epistemically rewarding” because it allows us to acquire information about the individual.

Although a mental file allows us to store information about that individual, the file’s reference is determined by the relation, not by the information. The referent is the individual the file is related to, independently of whether this individual satisfies all the information stored in the file. In Recanati’s terms, “The reference of a file is determined relationally, not satisfactionally” (2012, viii).

However, there can be mental files that are not related to any individual. Recanati calls them “irregular files”. He points out that “mental files are governed by an acquaintance norm” (2013, 208), which allows us to distinguish between regular and irregular files. A *regular* file is related to its object by an acquaintance relation, which is based on some suitable causal link or chain, viz. perception or testimony. Yet, there can be also *irregular* files lacking acquaintance. We normally deploy regular files, but we are also capable of deploying irregular files. In this sense, “one can think a singular thought in the absence of acquaintance” (Recanati 2012, 165). From the subject’s perspective, regular and irregular files behave roughly in the same way. They both give us the impression of being related to particular individuals in spite of the fact that the regular file really has an acquaintance relation to its object whereas the irregular file lacks it.

Within the category of irregular files, Recanati distinguishes two opposite cases. On the one hand, files that are created by a subject who attempts to satisfy the acquaintance norm by presupposing or expecting an acquaintance relation to an object, either correctly (as in the case of Neptune, an actual planet rightly predicted by Le Verrier) or incorrectly (as in the case of Vulcan, a nonexistent planet wrongly predicted by the same scientist). On the other hand, files that are created by a subject in spite of her awareness of the lack of an acquaintance relation. In the latter case, the subject does not attempt to satisfy the acquaintance norm and therefore she cannot be criticized for violating it. Still, the acquaintance norm remains operative, and is exploited by these files in order to produce their peculiar phenomenological effect: “the thought fails to have a singular content, though phenomenologically it feels as if it had a singular content” (Recanati 2013, 210). The latter is the relevant case for theorizing about fiction.

Mental files have been so far exploited in the debate about fiction mainly from an anti-realist perspective (cf. Crane 2013, Everett 2013, Friend 2014). The idea is that, in engaging with a work of fiction, one opens irregular mental files about fictional characters as if they were real people, albeit one knows that fictional characters in fact do not exist. My point is that this is not the whole story. Although I agree with the anti-realist idea that one opens irregular mental files about fictional characters as nonexistent individuals, I argue that one can also open regular mental files about fictional characters as existing abstract artifacts.

Unlike what happens in the case of ordinary objects, we can use two files, not just one, in our engagement with fictional characters: a file for a relation to a nonexistent individual, and another file for a relation to an existing abstract artifact. We, as appreciators of works of fiction, enjoy a purported relation to a nonexistent individual in virtue of exploiting the actual relation to an abstract artifact.¹ From this perspective, one can conceive of a fictional character as an abstract artifact whose function is to generate an irregular mental file about a nonexistent individual, and possibly also a regular mental file about the artifact itself. I will call the former a ‘fiction file’ and the latter a

'source file'. In the remainder of this section, I will focus on fiction files. I will discuss source files in the next section.

A fiction file purports to refer to a fictional character as a particular individual in the fictional world. It lacks an acquaintance relation to its object, and that is why it is irregular. Still, it differs from a mere figment of the imagination since it acquires information through an acquaintance relation between the subject and a real object – say, for the moment, a work of fiction. In this sense, a fiction file resembles a regular file. Nevertheless, the regular file acquires information through a relation to the object it is about whereas the fiction file gains information through a relation to another object, which is not the object the fiction file is about. In spite of its being irregular, the fiction file acquires information that is determined by public objective sources, not arbitrarily established by the thinker. In this sense, fiction files are not only irregular files but also “public or shared files – files shared by distinct individuals in a community” (Recanati 2012, 205). In our case, the community is the audience of a certain work of fiction. Each member of the audience opens her own fiction file, but is committed to store in this file the same core information stored by the other members of the audience.

2. A further step in the characterization of fiction files consists in treating them as special cases of those irregular files to which Recanati attributes a “discourse-referential function”. If we hear a speaker saying “I don’t believe that Mary had a baby and named *her* Sue”, we open not only a regular file about Mary (in addition to an already opened regular file about the speaker), but also an irregular file about Mary’s baby. According to Recanati, the reason is that “when we process discourse, we treat even non-referential pronouns *as if* they referred to objects, and store the information conveyed about the objects in question (‘the discourse referents’) in files, just as, in processing referential discourse, we store information conveyed about ordinary referents in files” (2012, 174).

A file that performs a discourse-referential function purports to refer within a framework embedded into a discourse-operator. In Recanati's example, the discourse-operator is made explicit by the clause "I don't believe that...", while in the case of fiction files the discourse-operator normally remains implicit, though it can be made explicit by clauses such as "according to the fiction..." or "in the fictional world...".

For a fiction file is an irregular file that purports to refer to a particular individual in the framework of a discourse, it can be underlain by a regular file that refers to the *source* of the discourse. That is why I call the latter file a 'source file'. Thus, information concerning fictional characters is split between two files. On the one hand, the *fiction file* stores "internal" (or "nuclear" or "fictional") information concerning the character as a particular individual, which I call the *f-character*. On the other hand, the *source file* stores "external" (or "extranuclear" or "meta-fictional") information concerning the character as something created, which I call the *s-character*. For example, the fiction file stores the date of birth of the f-character, the name of her parents, her job, while the source file stores the date of creation of the s-character, the name of her author, her role in the narrative structure (e.g. protagonist, antagonist, mentor). In sum, an s-character generates a fiction file, which purports to refer to the corresponding f-character, but we can actually refer to the s-character itself by means of a source file.

In fact, a fictional character can exist only as an s-character, but since the latter generates a fiction file, the audience of a work of fiction can experience also the corresponding f-character. When the appreciators of a work of fiction find a character especially significant, they can supplement the generic source file concerning the work of fiction with a specific source file concerning the fictional character itself. In this case, the external information about the s-character is unified into a specific source file instead of being grouped with the external information about other s-characters into the generic source file about the work of fiction. Creating a specific source file about the s-character, in addition to the fiction file concerning the f-character, suggests that we can think of a fictional character in two different ways. We can think of it as an individual who

inhabits the fictional world on the one hand, and as an artifact that contributes to the functioning of a work of fiction on the other.

Murray Smith (2011) characterizes such a duality as the “twofoldness of a fictional character”, noting that we can take two distinct attitudes towards fictional characters. On the one hand, “So readily do we recognize in fictions those ‘virtual persons’ we call characters that we can speak of them and respond to them in many ways just as if they were actual persons” (2011, 280). On the other hand, “We exhibit awareness of the configurational aspect of character whenever we note or notice something bearing upon the designed status of a character, when we see a character as an element in a representation” (2011, 280).

Developing Smith’s insight in terms of mental files, I call ‘twofileness’ the situation in which an appreciator of a work of fiction opens two files about a certain fictional character; namely, a fiction file about an individual in the fictional world, and a source file about an element in a representation. In this sense, twofileness allows the recipient of a work of fiction to experience a fictional character as a fictional individual (through the fiction file), while, at the same time, tracing this experience back to its source (through the source file). This is the main explanatory advantage of the notion of twofileness, that is, to provide us with a satisfying account of the complex cognitive activity of fiction appreciation.

In principle, there are three levels at which one can engage with fictional characters. At a basic level, one can just deploy fiction files together with some background awareness of the fact that such files do not refer to individuals in the actual world but only purport to refer to individuals in the fictional world. In this case, no source file is deployed even though the awareness of the nature of fiction involves acknowledging that there is some source that is responsible for the fictional story one is enjoying; therefore, in principle, one might refer to this source. That is to say that the source file is not actually deployed but there is a potential for deploying it. Arguably, this is the level at which very young children – and, more generally, “naive” or “immersed” recipients – engage with works of fiction.

At an intermediate level, one can start focusing on what is responsible for the information one has stored in the fiction files. This leads one to deploy not only fiction files about fictional individuals but also a source file about the work of fiction that provides one with information about such individuals. In this case, the external information about the s-character is stored in the file concerning the work of fiction, instead of in a specific source file. This can be considered as a basic case of twofileness in which a character shares its source file with other characters belonging to the same work of fiction. Arguably, this is the level at which ordinary appreciators usually engage with works of fiction.

Finally, at the highest level, one can gather information about the specific *design features* that are responsible for the information one gathers about a specific fictional individual.ⁱⁱ This leads one to deploy not only fiction files about fictional individuals and a source file about the work of fiction, but also source files about specific fictional characters understood as elements in a representation, that is, as functional components whose function precisely is the generation of fiction files.

From this perspective, a fictional character is, so to speak, the “source code” that leads one to generate a fiction file about the corresponding fictional individual, and a source file allows one to trace one’s experience of a fictional individual back to such a “source code”.ⁱⁱⁱ One can do that, for instance, by underlining all the sentences of a novel that contribute to one’s mental file about a certain fictional individual. This sort of exercise can improve our appreciation of that novel by leading us to pay attention to the “source code” through which the writer made up that fictional individual.

Arguably, this is the level at which sophisticated (“aesthetically-minded”) appreciators such as critics or scholars engage with works of fiction. However, this does *not* mean that fictional characters exist as functional components only for such sophisticated appreciators. Rather, the latter appreciators highlight something that was already there (namely, a “source code”, an element in a representation), but remained unnoticed in ordinary practices of appreciation. Sophisticated appreciators highlight the existence of fictional characters (understood as functional components) to

whom ordinary appreciators do not pay attention, just as the latter highlight the existence of the work of fiction to whom naive or immersed appreciators do not pay attention.

3. The notion of twofileness allows us to reconsider the debate between artifactualism and anti-realism about fictional characters from a new perspective. Let us begin with pointing out two basic premises on which both artifactualists and anti-realists can agree.^{iv}

First, a fiction file does not concern a real entity: we have the impression that a fictional individual exists only because we deploy a fiction file about it, i.e. a representation about it. In this sense, the fictional character is “representation-dependent”, as Tim Crane (2013, 68) puts it.

Second, this representation does not arise from nothing. The appreciators of a work of fiction share their representations about a certain fictional character in such a way that they can find a reasonable agreement about the basic features of that character. There is *something* objective that elicits such representations and warrants their convergence.

Still, what is this “something”? What is the source of the representations? Here is where the disagreement between artifactualism and anti-realism arises. On the one hand, the anti-realist denies that this source is the fictional character as an abstract artifact, claiming that this source is nothing but the work of fiction: “the authors create works of fiction: that is enough” (Crane 2013, 80). On the other hand, the artifactualist can point out that if one looks into the work of fiction, one finds passages that introduce a certain fictional character and ascribe features to it. It is these passages, not the work of fiction in general, that allow the appreciators to deploy and share representations about a certain fictional character, namely fiction files.

I argue that these passages have a distinctive unity that comes from their fulfilling a distinctive function, that is, eliciting a specific representation from the appreciators. Thus, these passages give a specific contribution to the functioning of the work of fiction. In this sense, the information one can gather in a source file about a fictional character is not generic information about the work of fiction, but specific information about that fictional character as a specific outcome of human

creativity. A source file of this sort differs from the corresponding fiction file precisely because the former *actually refers* to something in the actual world (something more specific than the work of fiction as a whole, namely the fictional character as a "source code"), whereas the latter only *purports to refer* to something in the fictional world (namely the fictional individual).

This argument in favor of artifactualism, which I will call "the argument from mental files", can be articulated as follows:

Premise (1): if the functioning of an artifact requires the fulfillment of a function that is more specific than the function of the whole artifact, there must be an element of the artifact that carries out that specific function, and this element, in turn, is an artifact.

Premise (2): works of fiction are artifacts.

Premise (3): the functioning of a work of fiction requires the fulfillment of a specific function, namely, the generation of a mental file about a certain fictional individual.

Conclusion (4): there are artifacts that are elements of works of fiction and carry out the function of generating mental files about fictional individuals.

Premise (1) is supported by the ontology of artifacts.^v For instance, the functioning of a wristwatch requires the fulfillment of the specific function of indicating minutes, and the element of the wristwatch that carries out that specific function, namely, the minute hand, in turn, is an artifact. Premise (2) is supported by the ontology of works of art. Premise (3) has been defended above.

These premises lead to the conclusion (4) in the following way. Because of (1) there must be an element of the work of fiction that fulfills the specific function considered in (3), and because of (2) this element in turn, is an artifact, which one can call a fictional character.

This conclusion involves that a fictional character is an artifact in virtue of being a functional component of a work of fiction; hence the fictional character is an element in a representation; hence it is representation-dependent. Yet, this dependence does not prevent the fictional character

from being an artifact. Indeed, the character is an artifact precisely in virtue of this dependence. In order to elucidate this point, let us reconsider Crane's claim that a fictional character is "representation-dependent" (2013, 68). In fact, there are three ways in which this predicate can be read.

First, a fictional character, considered as a fictional individual, is representation-dependent in the sense that it depends on a *mental* representation, namely a fiction file, which purports to refer to it. In this case, the representation involves a purported reference that lacks a referent, and thus being representation-dependent surely entails nonexistence.

Second, a fictional character, still considered as a fictional individual, is representation-dependent in the sense that it depends on a *public* representation, namely a work of fiction. Such a dependence involves a causal connection between the public representation, i.e. the work of fiction as a cause, and the mental representation, i.e. the fiction file (which purports to refer to the fictional individual) as an effect.^{vi}

However, this is not the whole story. There is a third way in which a fictional character is representation-dependent, viz. *as an element in a representation*. The relevant representation, here, is the work of fiction as a public representation, and the fictional character is no longer considered as a fictional individual but rather as a functional component of the work of fiction: what in the work of fiction is responsible for the generation of the fiction file about the corresponding fictional individual. Thus, the dependence of the character on this representation is a kind of part-whole relationship. This kind of dependence, unlike those considered above, fits with the existence of the character. For the work of fiction exists, its components also should exist. From this perspective, the relationship between a work of fiction and its fictional characters can be compared to the relationship between an organism and its organs, or a machine and its modules. Fictional characters contribute to the functioning of a work of fiction much as organs contribute to the functioning of an organism, or modules to the functioning of a machine.

In sum, a fictional character as an abstract artifact is a component of a work of fiction (i.e. a *public* representation) that has the function of generating a fiction file (i.e. a *mental* representation) that purports to refer to the corresponding fictional individual. By appreciating the design features of fictional characters, sophisticated appreciators pay attention not only to the work of fiction as a whole (i.e., to the representation as a whole) but also to the functioning of its components (i.e., to the elements in the representation); they can be compared to entomologists who pay attention not only to an insect as a whole but also to the functioning of its organs.

Just as insects' organs are not created by entomologists, fictional characters as functional components are not created by sophisticated appreciators, who, instead, exploit their expertise in order to notice these elements in a work of fiction and to make reference to them. In this sense, one might treat the argument from mental files as an argument to the best explanation of our practices of appreciation of works of fiction.^{vii}

4. The argument from mental files is compatible with two different accounts of fictional characters that I shall call “austere artifactualism” and “robust artifactualism”.

According to *austere artifactualism*, a fictional character is nothing but a component of a work of fiction that fulfills a specific function, just as the plumbing of a building or the respiratory system of an animal. It does not exist *in addition to* the work of fiction but only as an element of it. It is simply what, in a work of fiction, is responsible for our purported reference to the corresponding fictional individual. The appreciator of a work of fiction can keep track of this element by means of a source file just as she keeps track of the fictional individual by means of a fiction file.

If one remains in the framework of austere artifactualism, then stating that there is a work of fiction *and* its fictional characters is just another way of stating that there is nothing but a work of fiction. That is because the existence of fictional characters as functional components is part of the existence of the work of fiction, and therefore by stating the latter one is already stating the former. Thus, adding the conjunct “and its fictional characters” to the claim “there is a work of fiction” is as

redundant as adding the conjunct “and its respiratory system” to the claim “there is an animal”. That is to say that austere artifactualism does not involve refutation of Crane’s above mentioned anti-realist claim: “the authors create works of fiction: that is enough”.

This might lead one to wonder whether austere artifactualism really is a kind of realism or should rather count as a kind of anti-realism.^{viii} The answer to this question depends on whether one conceives of realism about fictional characters as a claim of their existence *in addition to* the work of fiction or just *within* the work of fiction. I will not take a stand on this. My point, here, is just that treating fictional characters as functional components of works of fiction contributes to explaining our appreciation of those works thereby shedding some light on the nature of these characters. To settle the question of whether fictional characters, as functional components of works of fiction, deserve to be called ‘real’ in a metaphysically robust sense goes beyond the scope of this paper.

Thus, I acknowledge that the view I am advocating might not be a fully-fledged realist account of fictional characters. It might be only a kind of functionalism according to which fictional characters are just functional components that contribute to the whole functioning of a work of fiction. However, even if this was the case, my conception of fictional characters as generators of mental files would remain slightly different from a sheer anti-realist conception according to which it makes no sense to conceive of fictional characters except as nonexistent individuals in the fictional world. In this sense, austere artifactualism – or functionalism, if one prefers – can be seen as the minimal step one can take to distance oneself from anti-realism. This is also the minimal sense in which one can treat a fictional character as an abstract artifact, that is, as a component fulfilling a specific function thereby contributing to the functioning of an encompassing abstract artifact.

If one wants to shift from such an austere functionalist view to an unquestionably realist view, one should endorse *robust artifactualism*, which conceives of a fictional character not just as the set of features of a work of fiction that provides us with information about the corresponding fictional

individual, but rather as a self-standing entity, which can correspond to such features but does not boil down to them.

A straightforward way to move from an austere artifactualism to a robust one consists in endorsing Voltolini's (2006, 223–246) argument. This states that the existence of a work of fiction entails the existence of its fictional characters because the identity conditions of the former need to make reference to the latter. From Voltolini's perspective, fictional characters are more than just elements of works of fiction; they are things to whom works of fiction must make reference in order to exist.^{ix}

Furthermore, drawing on Thomasson's (1999, 130; 2003, 220) analogy between fictional characters and juridical laws (both understood as abstract artifacts), one can develop a version of robust artifactualism that conceives of a fictional character as a sort of norm that has been implicitly established by the maker of a work of fiction. In terms of mental files, one might conceive of this norm as prescribing which information a suitable recipient should gather in the corresponding fiction file. In this sense, a fictional character is a *normative* abstract artifact, just as a juridical law is: they both prescribe what we ought to do. Yet, juridical laws are explicitly embodied in texts or codes containing articles, whereas fictional characters are implicitly embodied in the text of a work of fiction.

To sum up, both austere artifactualism and robust artifactualism are compatible with the conception of a fictional character as the element of a work of fiction that fulfills the function of generating a fiction file, that is, a mental file about a fictional individual. Yet, according to austere artifactualism, the fictional character is nothing but the set of design features of the work of fiction which prescribe what information one should gather in the corresponding fiction file. By contrast, according to robust artifactualism, such passages correspond to a set of prescriptions, which, in principle, might exist independently of one particular work of fiction, as suggested by the possibility for a fictional characters to appear in several works.

In order to elucidate the distinction between austere artifactualism and robust artifactualism, let us consider the way in which Nathan Salmon articulates his artifactualist account of fictional characters: “Think [...] of the various *roles* that a director might cast in a stage or screen production of a particular piece of fiction. Now think of the corresponding characters as the components of the fiction that *play* or *occupy* those roles in the fiction. It is no accident that one says of an actor in a dramatic production that he/she is playing a ‘part.’ The characters of a fiction—the occupants of roles in the fiction—are in some real sense *parts* of the fiction itself” (1998, 78-79).

On the one hand, by treating fictional characters as “the occupants of roles in the fiction” and thus “in some real sense *parts* of the fiction itself”, Salmon seems to endorse a kind of austere artifactualism.^x On the other hand, a supporter of robust realism would rather state that fictional characters are *the roles themselves*, which are embodied in the parts of the work of fiction. According to robust artifactualism, fictional characters exist as “parts” of the work of fiction in virtue of existing as “roles” that determine how such parts should be. As a “part” of a work of fiction, a fictional character allows us to store in our mental file about the corresponding fictional individual the features that are prescribed by this very character as a “role”.

Conversely, according to austere artifactualism, the prescriptions lie in the parts of the work of fiction itself. As Salmon puts it, “It is of the very essence of a fictional character to be depicted in the fiction as the person who takes part in such-and-such events, performs such-and-such actions, thinks such-and-such thoughts. Being so depicted is the character’s *raison d’être* [...] It is of the very essence of Shakespeare’s *Hamlet* that there is indeed an object that is Hamlet. I am not urging that we countenance a person who is Hamlet₁ and who contemplated suicide according to the classic play but who does not exist. There is no sense in which there is any such person. The objective Hamlet is Hamlet₂—what plays the title role in the Bard’s drama—and hence not a human being at all but a part of fiction, merely depicted there as anguished and suicidal” (1998, 79).

With respect to Salmon’s example, twofileness consists in opening not only an irregular *fiction file* about the fictional individual Hamlet₁ but also a regular *source file* about Hamlet₂, i.e. the

element of the work of fiction that is responsible for our deploying the irregular file about Hamlet₁. However, the argument from mental files should remain neutral on whether Hamlet₂ is nothing but a part of the work of fiction *Hamlet* (as stated by austere artifactualism) or a self-standing role that is embodied in that part of *Hamlet* (as stated by robust artifactualism). The twofileness phenomenon is compatible with both these explanations.

5. There is a sense in which the information we store in the fiction file can be also contained in the source file. For instance, the piece of information that Hamlet is anguished can be stored in both files. Yet, the way in which this piece of information is stored in the two files is different since the two files concern different kinds of things. In the fiction file, ‘anguished’ is a feature of a person, whereas in the source file this is a feature of an artifact, which generates our fiction file and prescribes us to store this feature in it. In other words, in the fiction file ‘anguished’ indicates what a person is, whereas in the source file ‘anguished’ indicates what an artifact does; the person *is* anguished whereas the artifact *represents* that person as anguished. Borrowing Zalta’s (1983) terms, the property ‘anguished’ is “exemplified” by a person but only “encoded” by an abstract artifact, which instead exemplifies properties such as ‘having been created by Shakespeare’.^{xi}

The fact that the same features can be present in both files, albeit in different ways, allows us to exploit one file as a surrogate for the other. In particular, we can exploit the fiction file as a surrogate for the source file. When we do so, we grasp the features that the fictional character, as an artifact, actually prescribes by attending to the features that the fictional character, as an individual, fictionally possesses. In this sense, twofileness can be relevant not only for our engagement with fiction “from the inside”, but also for our discourse about fiction “from the outside”.

On the one hand, in our engagement with fiction, we normally focus on the fictional character as an individual to which the fiction file purports to refer, and the source file becomes relevant only when we wonder what caused our experience of that fictional individual. On the other hand, in our discourse about fiction, we rather focus on the fictional character as an abstract artifact to which our

source file actually refers, and the fiction file becomes relevant only when, in order to figure out what the artifact prescribes us to do, we abide by this prescription thereby deploying the corresponding fiction file. In this case, we figure out the cause (the artifact) by undergoing its effect (the fiction file).

This allows us to shed some light on the “mixed perspective” (Everett, 2013, 163) that one adopts when one thinks, for example, of Anna Karenina as *a fictional character who commits suicide*. In principle, when one thinks of fictional characters without directly engaging with a work of fiction, one can deploy only a source file containing both historical information (e.g. that Anna was created by Tolstoy) and design features (e.g. that Anna is represented as committing suicide). However, the most effective way we have to represent design features consists in deploying the fiction file that satisfies such prescriptions. In this sense, one might say that the source file contains *a pointer* to the corresponding fiction file. Therefore, thinking of Anna as a fictional character who commits suicide amounts to exploiting this pointer, which connects the source file about *a fictional character*, which represents a woman as committing suicide, to the fiction file about a woman *who committed suicide*. This suggests that twofileness can be relevant not only in our engagement with fiction but also in our discourse about fiction.

To sum up, thinking of Anna Karenina as a fictional character who commits suicide consists in putting in the *source* file concerning this character the information that it represents a woman named Anna as committing suicide, thereby being entitled to put in the corresponding *fiction* file the information that Anna committed suicide.^{xii} This seems to be a reasonable way of thinking of fictional characters from outside of the engagement with a fiction. Yet, this only holds for external discourse. In fact, when one actually engages with Tolstoy’s novel, it would be awkward to think of Anna as a ‘fictional character who commits suicide’. Indeed, *Anna Karenina* mandates the reader to think of Anna as *a woman* who commits suicide. That is to say that, *within the fictional world*, Anna does *not* have the feature of being represented by Tolstoy as committing suicide. She commits suicide, full stop. That is why we need a fiction file distinct from the source file.

In fact, the emotions that a novel such *Anna Karenina* can elicit from the reader come from thinking of Anna as a flesh-and-blood individual. Surely, these emotions can be tempered by simultaneously deploying the corresponding source file that reminds us that Anna is nothing but an abstract artifact. Yet, for these emotions to arise, a fiction file concerning Anna as *a woman* should be deployed.

6. One may wonder whether the argument from mental files can be extended also to works of non-fiction. Consider a biography of Napoleon. It may be tempting to say that, in reading it, we open a “fiction file” (or, better, a “history file”) in order to store information about Napoleon as an individual and also a “source file” in order to store information about Napoleon as a character in this book. For example, we store in the source file about Napoleon pieces of information such as “being the protagonist of this book” or “being introduced at page 3”. Yet, in this case, the argument from mental files would lead us to the conclusion that historical figures are, in turn, abstract artifacts. If one took this to be clearly absurd, one would conclude that there must be something wrong in the argument from mental files, something that should prevent it from showing that fictional characters are abstract artifacts.

However, I argue, there is nothing absurd in treating historical characters as components of works of history. In fact, these works function just as works of fiction do, that is, by generating mental files about individuals. Just as we conceive of fictional characters as the functional components of works of fiction that generate mental files about fictional individuals, we might conceive of historical characters as the functional components of works of history that generate mental files about historical individuals. Thus, in appreciating a work of history, we can focus on the way in which historical characters function as elements in a representation as well as we can focus on fictional characters as functional components of a work of fiction.

Still, there is a crucial difference between fiction and history (and, more generally, non-fiction) in this respect. On the one hand, the function of a historical character, as an element of a work of

history, consists in generating a *regular* file about an individual who belongs to the same world to which the work of history belongs. On the other hand, the function of a fictional character, as an element of a work of fiction, consists in generating an *irregular* file about a fictional individual who belongs to a fictional world that is different from that to which the work of fiction belongs.

As a consequence of this, the deployment of mental files in our engagement with works of history is different from that in our engagement with works of fiction. For instance, in reading a biography of Napoleon, pieces of information such as “being the protagonist of this book” or “being introduced at page 3” can be stored directly in the file that we deploy about Napoleon as a historical individual. This is because the biography we are reading belongs to the same world to which Napoleon belongs, namely the actual world, and, within this world, there is a causal connection that relates Napoleon to his biography, thereby allowing us to treat some features of that biography as features of Napoleon himself. Thus, storing such pieces of information as “being the protagonist of this book” or “being introduced at page 3” in the mental file about Napoleon as an individual is a correct way of reading a biography of Napoleon.

By contrast, if we consider a novel about Pinocchio, “being the protagonist of this book” or “being introduced at page 3” are not features that a suitable reader should ascribe to Pinocchio as an individual. These are just features of Pinocchio as an abstract artifact, not features of Pinocchio as the inhabitant of a fictional world. The reason is that the novel we are reading is not part of the fictional world to which Pinocchio as an individual belongs, and therefore the pieces of information about the novel cannot be ascribed also to a fictional individual.

In this sense, the distinction between fiction files and source files allows us to highlight a distinction between works of fiction and works of non-fiction. For what concerns pieces of information such as “being the protagonist of this book” or “being introduced at page 3”, a work of non-fiction allows us to store them in the same mental file in which we store information about the protagonist of the work as an individual. By contrast, a work of fiction mandates us to store such pieces of information in a distinct mental file. That is because a work of fiction and its fictional

individuals belong to distinct worlds (viz. the actual world and the fictional world, respectively), and therefore the pieces of information that concern the work of fiction cannot concern also its fictional individuals. Otherwise, in order to warrant the appropriate flow of information, we should treat the fictional world as causally connected to our actual world, in direct contrast with what works of fiction typically mandate us to do.

7. So far I have considered mental files deployed by the users of works of fiction, and I have characterized fictional characters as artifacts whose function consists in the generation of mental files. More specifically, the fictional character as an abstract artifact generates a fiction file in the appreciator's mind that purports to refer to the corresponding fictional individual. Furthermore, the appreciator can also open a source file that actually refers to the fictional character as an abstract artifact.

At this point, one might wonder how an abstract object such as a work of fiction, or a fictional character, can originate a causal chain that ends up in the generation of mental files in an appreciator's mind. On the one hand, I acknowledge that the proximal cause of the generation of mental files in the appreciator's mind is *a token* of the work of fiction, e.g. a paper copy of a novel or a screening of a movie. On the other hand, I argue that this token is only the penultimate link of a causal chain that originates in the creation of the work of fiction – and of fictional characters as its components – and ends up in the generation of mental files in the appreciator's mind. From this perspective, authors create abstract objects that, within the relevant cultural practice, establish the way in which the corresponding tokens are to be produced and used.^{xiii} Thus, we can deploy mental files about such abstract objects by exploiting a direct causal relation to their tokens.

This is what usually happens when we appreciate novels or movies. We exploit a paper copy or a screening in order to acquire the relevant information, but the mental files in which we gather information about these works usually are about novels as abstract objects or movies as abstract

objects, not about their concrete tokens. That is why appreciators can refer to the *same* work of fiction in discussion even though they have been in touch with different tokens of this work.

That being the case, if we treat fictional characters as components of works of fiction such as novels and movies, we can conclude that an appreciator is causally connected to fictional characters as abstract objects in the same way in which she is causally connected to works of fiction as abstract objects.^{xiv} Concrete tokens of a work of fiction are links of a causal chain that originates in the creation of the work of fiction itself as an abstract artifact, as well as of the fictional characters that are its components, and ends up in the generation of mental files in the appreciator's mind.

Certainly, an author does not have to possess the notion of a mental file, yet mental files provide us with an effective model of the process through which a character is created. For instance, Honoré de Balzac designed Lucien de Rubempré in such a way that a suitable reader of *Les illusions perdues* should store in a fiction file internal information like that f-Lucien is a writer, and can store in the corresponding source file external information like that s-Lucien is the hero of the novel and one of the main characters of the narrative cycle called "La Comédie humaine".

The main task of the creator of a fictional character consists in providing the audience with the internal information to be stored in the fiction file. The external information to be stored in the source file often remains implicit, as a side effect of the creation of the work of fiction within a certain cultural practice. However, an appreciator can take such information into account in order to properly appreciate the design of a certain character. The role of the source file becomes more explicit when authors do not create fictional characters from scratch but build up new versions of already existing characters. In this case a suitable appreciator of this work of fiction should store this information concerning the character's history of making in the source file about that character.

In the process of creation of fictional characters, authors can – and arguably do – deploy irregular mental files about fictional individuals. Yet, such "pre-fiction" files are not to be confused with the fiction files that will be generated in the appreciator's mind. Although both pre-fiction files and fiction files are irregular files, the former are mere figments of the author's imagination

whereas the latter are generated (and constrained) by public sources, that is, abstract objects that are causally active through concrete tokens such as paper copies of novels or screenings of movies. That is to say that pre-fiction files and the fiction files have a different role in the life cycle of a work of fiction; the former can contribute to its creation whereas the latter are a result of its functioning. In sum, an author creates fictional characters through a process that may involve the deployment of *pre-fiction* files, but the outcome of this process are fictional characters as abstract artifacts, which should be capable of generating *fiction* files, as well as source files, in the appreciator's mind.

For I conceive of a fictional character as an artifact whose primary function consists in generating fiction files thereby contributing to the functioning of a work of fiction, I state that the creation of a fictional character is completed when the creation of the corresponding work of fiction is completed. This is my way of addressing the “when-problem” that Stuart Brock (2010, 355) formulates in the following terms: “One problem for the abstract creationist, then, is specifying *when* fictional characters are brought into existence by their author. When, exactly, does life begin for a fictional character?”. My functionalist approach involves that fictional characters start existing *when* they can properly contribute to the work of fiction of which they are functional components, that is, *when* the work of fiction in which they appear starts existing. Until a work of fiction is nothing but a work in progress, its fictional characters are in turn nothing but “artifacts in progress”. The life of a fictional character begins when the corresponding work of fiction begins existing, just as, for example, the black king in chess begun existing when the game of chess did so.

8. Ultimately, a fictional character is an abstract artifact, which is created by an author (or a team thereof) by creating a work of fiction and whose function is to generate a fiction file in the mind of a suitable appreciator of that work of fiction, and possibly a source file that allows one to also appreciate the design features of that character. Although *all* fictional characters, as abstract artifacts, have the function of generating a fiction file and possibly a source file, the nature of the

information stored in these files depends on the medium through which a certain character is created. For example, in literary fiction the information stored in the fiction file basically consists of linguistic predicates, whereas in staged fiction (e.g. theater, cinema, TV series) the fiction file can store information in a sensory format. Furthermore, in the case of staged fiction, the source file stores not only information directly supplied by the creator of the fictional character but also information about the staging, for example information about the actor who plays the character.

We, as readers of novels or spectators of movies, carry out tasks of identification and recognition of particular individuals that emulate the corresponding tasks which our ordinary experience relies on. For example, the task carried out by the spectator of a movie emulates her ordinary task of perceptual identification and recognition of particular individuals, and the task carried out by the reader of a novel emulates her ordinary task of identification and recognition of particular individuals referred to in oral or written communication. In our engagement with works of fiction, we can effectively carry out such tasks because fiction files have a phenomenology of reference that is similar to that of regular files.

The key difference is twofileness, that is, the fact that we can supplement the fiction file with a source file that discloses the true nature of the fictional individual to whom that file purports to refer. On the one hand, by deploying a regular file in our *ordinary experience*, we have the impression that we are referring to a particular individual and we are in fact referring to her. On the other hand, by deploying a fiction file in our *experience of works of fiction*, we have the impression that we are referring to a particular individual, when in fact the thing to which we are actually related to is an abstract artifact, and we can pay attention to this by deploying a source file that actually refers to this artifact. While the fiction file allows us to enjoy a fictional character as if it was a real person, the source file reminds us of what it really is.^{xv}

Bibliography

- Brentano, Franz, 1982, *Descriptive Psychology*. Edited and translated by B. Müller. London and New York: Routledge, 1995.
- Brock, Stuart, 2010, “The Creationist Fiction: The Case against Creationism about Fictional Characters”, *Philosophical Review*, 119, 3, 337–364.
- Crane, Tim, 2013, *The Objects of Thought*, Oxford: Oxford University Press.
- Dennett, Daniel, 1987, *The Intentional Stance*, Cambridge, MA: MIT Press.
- Dodd, Julian, 2007, *Works of Music: An Essay in Ontology*, Oxford: Oxford University Press.
- Everett, Anthony, 2013, *The Nonexistent*, Oxford: Oxford University Press.
- Friend, Stacie, 2014, “Notions of Nothing”, in M. Garcia-Carpintero and G. Marti (eds.) *Empty Representations*, Oxford: Oxford University Press, 307–332.
- Giustina, Anna, 2017, ‘Conscious Unity from the Top Down: A Brentanian Approach’, *The Monist*, 100(1): 15–36.
- Grice, Paul, 1969, “Vacuous Names”, in D. Davidson and J. Hintikka (eds.) *Words and Objections*, Dordrecht: Reidel: 118–145.
- Kriegel, Uriah, 2018, *Brentano’s Philosophical System: Mind, Being, Value*, Oxford: Oxford University Press.
- Kroes, Peter, 2012, *Technical Artefacts: Creations of Mind and Matter. A Philosophy of Engineering Design*, Dordrecht: Springer.
- Lopes, Dominic, 2005, *Sight and Sensibility: Evaluating Pictures*, Oxford: Clarendon Press.
- Recanati, François, 2012, *Mental Files*, Oxford: Oxford University Press.
- , 2013, “Mental Files: Replies to my Critics”, *Disputatio* 36: 205–240.

- Salmon, Nathan, 1998, “Nonexistence”, reprinted in *Metaphysics, Mathematics, and Meaning – Philosophical Papers I*, Oxford: Oxford University Press, 2005, 50–90.
- Smith, Murray, 2011, “On the Twofoldness of Character”, *New Literary History* 42.2: 277–294.
- Soames, Scott, 2002, *Beyond Rigidity*, Oxford: Oxford University Press.
- Strawson, Peter, 1959, *Individuals*, London: Methuen.
- , 1974, *Subject and Predicate in Logic and Grammar*, London: Methuen.
- Thomasson, Amie, 1999, *Fiction and Metaphysics*, Cambridge: Cambridge University Press.
- , 2003, “Speaking of Fictional Characters”, *dialectica* 57, 2, 205–223.
- Van Inwagen, Peter, 1977, “Creatures of Fiction”, *American Philosophical Quarterly*, 14: 299–308.
- Voltolini, Alberto, 2006, *How Ficta Follow Fiction*, Dordrecht: Springer.
- Zalta, Edward, 1983, *Abstract Objects: An Introduction to Axiomatic Metaphysics*, Dordrecht: Reidel.

i

This does not entail that, in engaging with fiction, one pretends, *of* an abstract artifact, that this is a flesh-and-blood individual. The account I am advocating is not committed to this claim and therefore can remain neutral on this issue (a thorough discussion of which can be found in Thomasson 2003, 210–214).

ii

I borrow the notion of ‘design features’ from Dominic Lopes’ account of depiction. According to Lopes (2005, 25), *design features* are those features of the surface of a picture that allow a suitable viewer to enjoy a visual experience of a certain object depicted. Likewise, in my conception of fictional characters, *design features* are those features of a work of fiction that allow a suitable appreciator to deploy a mental file that purports to refer to a certain fictional individual. In the wake of Kroes’s account of technical artifacts, I conceive of the design of an artifact as the way in which “a function is ‘translated’ into a structure” (2012, 10). In the case of fictional characters, the function consists in the generation of mental files while the structure is constituted by design features.

iii

In the case of video games, where a fictional character is encoded in a module of a computer program, the source-code metaphor can be taken literally.

iv

What I am considering, here, is a version of anti-realism like Crane’s, according to which only works of fiction exist, not fictional characters. I am not considering more radical anti-realist views according to which works of fictions do not exist either.

v

Consider the ontological principle according to which “Technical kinds are functional kinds; being an instance of a technical kind is determined by having the appropriate proper function” (Kroes 2012, 98). My point is that the “appropriate proper function” that consists in generating a mental file about a fictional individual within the encompassing functioning of a work of fiction individuates the fictional character as a “technical kind”. Consider also that a fictional character, just as a technical artifact, can malfunction. This happens, for instance, when it leads one to store inconsistent information in the mental file about the corresponding fictional individual.

vi

I am assuming here that public representations such as works of fiction can have causal efficacy. I will defend this claim in § 7.

vii

Thanks to an anonymous reviewer for pointing out the possibility to interpret my argument as an argument to the best explanation.

viii

Thanks to both the anonymous reviewers for leading me to address this issue.

ix

Another way of expressing the difference between austere and robust artifactualism consists in considering the dependence relationship between the fictional characters as parts and the work of fiction as a whole. According to austere artifactualism, the parts depend on the whole: if the whole had not existed, the parts would not have existed since their existence consist in contributing to the functioning of the whole. According to robust artifactualism, it is the other way around: if the parts had not existed, the whole would not have existed since its identity conditions require making reference to such parts. From this perspective, one might treat austere artifactualism as a sort of instrumentalism that cast fictional characters as effective explanatory posits or “illata” (along the lines of Dennett 1987), or also as parts that are conceptually distinguishable even though they are not actually separable (along the lines of Brentano 1982/1995, 16; see also Giustina 2017, 28–30; and Kriegel 2018, 174–178). Thanks to an anonymous reviewer for leading me to consider this way of contrasting austere and robust artifactualism.

x

So does Scott Soames (2002, 93): “Among the things that exist are stories, legends, novels, chapters, plays, movies, and the like. These are abstract objects created by authors. Fictional characters are constituents of these objects. Like the fictions of which they are a part, fictional characters are a special kind of real, existing object”.

xi

Van Inwagen (1977, 305) makes a similar distinction between proprieties that are exemplified by a fictional character and features that are only “ascribed” to it.

xii

With respect to Zalta’s distinction introduced above in this section, I state that Anna as an abstract artifact *exemplifies* the feature “representing as committing suicide” whereas it only *encodes* the feature

“committing suicide”. This enables one to store in the source file about Anna the information that it is a fictional character that *in the story* commits suicide (cf. Thomasson 1999, 105). I am grateful to an anonymous reviewer for leading me to articulate this point.

xiii

An author can create a work of fiction as an abstract object by producing what Strawson (1959, 232) calls “a model particular”, as for example the manuscript of a novel: “a kind of prototype, or ideal example, itself particular, which serves as a rule or standard for the production of others”. However, the ultimate outcome of the author’s creation is the work of fiction as an abstract object, not as a model particular, as shown by the case in which the model particular is lost or destroyed and yet the work of fiction keeps existing.

xiv

A thorough account of the causal power of abstract objects with which I am rather sympathetic (except for its Platonism), is provided by Julian Dodd (2007). Consider for instance this passage: “it only *seems* obvious that abstract (i.e. non-spatial) entities cannot be causally efficacious, once a couple of controversial philosophical assumptions have been made. When these assumptions are (rightly) questioned, the idea that certain *abstracta* may yet be causally active [...] turns out to be harmless [...] Now, the assumptions made by someone who denies that abstract entities such as types can be causally efficacious would seem to be these: first, that there are clear criteria for whether an entity participates causally in an event; and, second, that it is clear that entities without a spatial location fail to meet these criteria. But in fact, neither assumption is warranted” (Dodd 2007, 13). For what concerns the possibility to actually refer to abstract objects through chains of communication, see the chapter “Reference to fictional characters” in Thomasson’s *Fiction and Metaphysics* (1999, 43-54). Consider, for instance, this passage: “Indeed the basic model that Kripke offers in *Naming and Necessity* and Gareth Evans modifies in *The Varieties of Reference* of the determination of the reference of names through a “baptism” and the continuation and proliferation of the use of a name via chains of communication can be applied in a modified version to names of fictional entities — provided we allow that not only causal relations but also relations of ontological dependence can form a path for direct reference” (Thomasson 1999, 44). Thanks to a reviewer for leading me to focus on how fictional characters can originate causal chains that end up in the generation of mental files in the minds of appreciators.

xv

I want to thank Filippo Contesi, Adrien Glauser, Enrico Grosso, François Recanati, Louis Rouillé, Daniela Tagliafico, Alberto Voltolini and the anonymous referees for their insightful comments on previous versions of this paper. I also want to thank the organizers and the audience of the Mind and Language Seminar of the Institut Jean Nicod, Paris, where I presented the first version of this paper on October 7th 2015.