

## Against Sainsbury's Irrealism About Fictional Characters: Harry Potter as an Abstract Artifact

Our commonsense ontology includes ordinary objects like cups, saucers, tea kettles and jars of honey; buildings like the Big Ben; people alive today, like the novelist J.K. Rowling. We consider them part of our reality. We are commonsense realists about objects like cups, saucers, the Big Ben and J.K. Rowling. On the face of it, this is a fairly strong reason to include such objects in the ontology we posit. What about fictional characters like Harry Potter? Our intuitions pull us in opposite directions. On the one hand, clearly, Harry Potter—unlike the Big Ben—doesn't exist, for if we inventoried the objects and people existing today (or that ever existed), Harry Potter wouldn't be among them. On the other hand, we want to say the character didn't exist prior to the 1990s, before J. K. Rowling thought up the novels, but through her imagination and authorial intent to create a fictional world filled with fictional heroes and villains, she brought Harry, Dumbledore and a host of other fictional characters into *existence*. In other words, at first blush, intuitions about authorial creation support including fictional characters in our reality, while our commonsense ontology speaks against realism about fictional characters.

In this paper, I will explore arguments for and against one form of realism about fictional characters: *abstract artifact theory* about fictional characters ('*artifactualism*' for short), the view according to which fictional characters are part of our reality, but (unlike concrete entities like the Big Ben and J. K. Rowling), they are abstract objects created by humans, akin to the institution of marriage and the game of soccer. I shall defend this view against an objection that Mark Sainsbury (2010) considers decisive against artifactualism: "When we think about fictional entities, we do not think of them as abstract. Authors, who ought to know, would fiercely resist the suggestion that they are abstract. Abstract artifact theory entails that producers and consumers of fiction are sunk in error" (111). In other words, artifactualism attributes to people who produce and process sentences and thoughts about Harry Potter massive error, indeed, a category mistake about what kind of thing Harry Potter is. For an abstract object (such as the institution of marriage) isn't the sort of thing that can wear

glasses, ride a double-decker bus, attend school. I shall call this the *category-mistake objection*.

In Section 1, I will distinguish artifactualism from various other forms of realism about fictional characters, and from the position of *irrealism* about the likes of Harry Potter, a view according to which fictional characters don't exist; only the works of fiction portraying them do. In Section 2, I will explore one general and powerful argument for favoring artifactualism over other realist alternatives: it can successfully account for authors creating fictional characters. In Section 3, I will consider and deflect the category-mistake objection, which, according to Sainsbury, gives an edge to irrealism over artifactualism. Artifactualism, I shall conclude (in Section 4), remains a tenable contender.

## 1. REALIST AND IRREALIST POSITIONS ABOUT HARRY POTTER

We may, along with Mark Sainsbury (2010, 44–114), distinguish three realist alternatives about fictional characters: there really are such things just as there are ordinary concrete objects occupying space and time; but unlike those ordinary objects like cups, saucers and the Big Ben, ...

- fictional characters *don't exist*, according to *Meinongianism* about fictional characters;<sup>1</sup>
- fictional characters *are not actual* but merely possible, according to *nonactualism*;<sup>2</sup> and
- fictional characters *are not concrete* but abstract, created by the activities of authors according to *artifactualism*.<sup>3</sup>

<sup>1</sup> For brevity's sake, I'll suppress the qualification 'about fictional characters' and will simply talk of realism, irrealism, Meinongianism, nonactualism, artifactualism, Platonism. Whenever these labels appear unqualified, they are shorthand for theories about fictional characters.

Parsons (1980) is a contemporary proponent of Alexius Meinong's (1904) eponymous theory.

<sup>2</sup> Lewis (1978) put forth such a view. This position is sometimes called possibilism about fictional characters. See also Kripke's earlier (1963) view about Sherlock Holmes.

<sup>3</sup> Kripke (1973), Searle (1974/1979), van Inwagen (1977), Fine (1982), Schiffer (1996), Salmon (1998), Thomasson (1999) are prominent proponents who hold that authors' creative process of writing novels, stories, etc. creates fictional characters. This position is sometimes called creationism about fictional characters.

There is a position in logical space for holding that fictional characters are abstract but exist timelessly, and authors don't create but discover them—we might call such a view Platonism about fictional characters. Zalta's (1983) unorthodox neo-*Meinongian* proposal can be considered an instance of such an account. The only kind of abstract-object theory I will consider in this paper is artifactualism, given the overwhelming popularity and attention that this position has been enjoying (compared to Platonism), as well as the advantages that I think it has over rival theories (Platonism included) precisely because it treats fictional characters as human-created objects.

One of the chief motivations for Meinongianism is this: plausibly, there are many things that don't exist, things that, while nonexistent, are the objects of our thought and imagination. Harry Potter is one of them, as is the batch of vanilla pudding I considered cooking up this afternoon (from a specific packet of pudding and batch of milk and sugar), but never got around to it. A similar motivation drives the nonactualist position: the range of things that are possible extends beyond things that are actual: the pudding I might have cooked this afternoon is a nonactual, merely possible object; as is Harry Potter.

The artifactualist position raises the intricate issue of deciding what exactly the abstract/concrete distinction consists in. The assumption so far has been that abstract objects (unlike concrete ones) don't occupy space and time. Another option is that abstract objects (unlike concrete ones) lack causal powers. A third option is to identify paradigmatic examples of concrete and of abstract objects in order to illuminate the distinction.<sup>4</sup> I won't dwell on these options here, because the ways in which the abstract/concrete distinction is traditionally drawn are called into question precisely in the light of abstract object created by human activity, *abstract artifacts*, that is—for example, the institution of marriage and the game of chess.<sup>5</sup>

Instead of defining the categories of abstract versus concrete, I will therefore take as my point of departure a broad and fairly uncontroversial range of examples for both concrete and abstract objects. Concrete objects clearly include things like cups, saucers, actual batches of pudding, the Big Ben, J. K. Rowling. Many of those who posit abstract objects count among them numbers, sets, propositions and properties like being tall and being human. Those who consider these abstract objects agree that typically,<sup>6</sup> they are *timelessly existing abstract objects that are mind-independent* in the following sense: their existence at a time *t* is independent of any mental activity at *t*.

<sup>4</sup> For an overview of these and other ways of drawing the abstract/concrete distinction, see Rosen (2012).

<sup>5</sup> See Rosen (2012) and Fine (1982, 130–131) motivating the claim that there should be room for abstract objects that come into existence contingently:

...what underlies the platonist's position is a certain ontological prejudice. ... These philosophers suppose ... that certain features should go together, so that the same entities will be material, will exist in space and time, will exist contingently, etc., and the same entities will be immaterial, not exist in space and time, be necessarily existent, etc. Now although paradigmatic cases of concrete and abstract objects may have exactly the features from one or other of these groups, it must be recognized that there are objects of intermediate status that share features from both.

<sup>6</sup> Barring exceptions like the singleton set of my red mental image upon spotting a strawberry. There are various ways to go on the status of such sets: we could conclude that not all sets are abstract after all or that the notion of mind-independence at work should be revised to allow such sets to be abstract (see Rosen 2012).

There is, however, another type of abstract object one might posit: *abstract artifacts*. Notice that an abstract artifact like the game of chess does have temporal features, after all: the game of chess didn't exist before 1000 A.D. and has been in existence for several centuries (but beyond that, there is disagreement). Nonetheless, an abstract artifact would still be mind-independent in the above sense: the game of chess can exist at a time without anyone having any chess-related mental activity at that time. It's worth giving a variety of examples of abstract artifacts:

- the games of soccer and chess; the chess move of castling;
- the institution of marriage and the office of prime minister;
- religions like Anglicanism or Buddhism;
- traditions like Mardi Gras and Easter celebrations;
- more specific traditions like the New Orleans Mardi Gras festival and the Village Halloween Parade in New York City;
- the tradition of the Easter bunny and its Australian marsupial counterpart, the Easter bilby;
- recipes for preparing treacle pudding and beef Wellington;
- the letters of the alphabet ('A', 'Z', etc.);
- brands like Twinings, Twix and Mini Cooper;
- words and names of a language, including fairly recent additions like 'netiquette' (rules governing polite behavior in interactions on the internet) and 'cot potato' (for a small child spending a lot of time in front of the TV set, that is, a very young couch potato); also the first name 'Dweezil' for boys, coined by Frank Zappa;
- musical works like Mozart's serenade *A Little Night Music* and opera *The Magic Flute*;
- literary works like the seven Harry Potter novels written by J. K. Rowling.

We thus have a long and varied list of candidates for abstract social and cultural (legal, artistic, religious, linguistic etc.) artifacts among which it is natural to make room for fictional characters like Harry Potter also—the defender of artifactualism suggests.

Alternatively, a fourth option is to forgo realism about fictional characters, opting for *irrealism*, which denies all forms of ontological commitment to fictional entities.<sup>7</sup> Irrealism proposes to analyze sentences like (1)-(8) below with the help of an ontology that is committed only to the existence of works of fiction: novels, films, and so on:

<sup>7</sup> Sainsbury (2010) favors this view, as does Walton (1990).

- (1) Harry Potter doesn't exist.
- (2) "Harry had a thin face, knobby knees, black hair, bright green eyes. He wore round glasses held together with a lot of Scotch tape..."<sup>8</sup>
- (3) "From his first days at Hogwarts, the young, green-eyed boy bore the burden of his destiny as a leader, coping with the expectations and duties of his role..."<sup>9</sup>
- (4) Harry Potter is a fictional character.
- (5) Harry Potter was created by J.K. Rowling.
- (6) Stephen King thinks about Harry Potter.<sup>10</sup>
- (7) Harry Potter is more famous than Sparrowhawk (the wizard of Earthsea).<sup>11</sup>
- (8) Harry Potter is not as reckless as Sparrowhawk.

(I have included (1) here in order to have a complete list of the types of sentences that the various theories about fictional characters have to account for. Of course, capturing the truth of (1) is easy for an irrealist and tricky for realists. I won't discuss realists' proposals for analyzing (1) in this paper.)

(2)-(8), when taken at face value, seem to accrue ontological commitment to characters from fiction, and therefore provide *prima facie* evidence for one or another of the realist positions. It is well to note that (2), (3) and (8) form a separate group on this list: it isn't *literally* true that Harry Potter had knobby knees, wore glasses etc.; it's true according to the Harry Potter fiction, but if we enumerate all those who wear glasses, Stephen King and the rest, Harry wouldn't be on the list. (2) is quoted from one of the Harry Potter novels. (3) discusses the content of another Harry Potter novel; (8) compares the content of two works of fiction; we can call (2), (3) and (8) *fiction-internal* sentences put forth by authors, readers (including critics). As before, we can say that it isn't *literally* true that Harry Potter is a young green-eyed boy who bore the burden of his destiny as a leader. And it isn't literally true that Sparrowhawk is more reckless than Potter. It is therefore unclear just how great a challenge (2), (3) and (8) pose for irrealism. The irrealist might argue that though each is literally false, they can be replaced, respectively, by the true (2'), (3') and (8'), hence their true ring:

<sup>8</sup> From Volume 1: *Harry Potter and the Sorcerer's Stone*, page 20.

<sup>9</sup> The literary critic Michiko Kakutami wrote this in her review of volume 7 in *The New York Times*, "For Harry Potter, Good Old-fashioned Closure", 07/04/2007.

<sup>10</sup> Stephen King has been a major fan of Harry Potter and a defender of the Harry Potter series against criticism from literary theorists.

<sup>11</sup> Ursula K. le Guin began writing her now-classic series of fantasy novels and short stories about Earthsea in the 1960s; they feature Sparrowhawk, a young orphaned boy who discovers he has magic powers.

- (2') According to the first Harry Potter novel, Potter had a thin face, knobbly knees, black hair, bright green eyes, wore round glasses held together with Scotch tape.
- (3') According to the seven Harry Potter novels, Harry Potter was a young, green-eyed boy who, from his first days at Hogwarts, bore the burden of his destiny as a leader, coping with the expectations and duties of his role.
- (8') According to the Harry Potter and Earthsea fiction series, Potter is not as reckless as Sparrowhawk.

(4)–(7) are more complicated, however: they aren't just true according to a body of fiction; they appear to be literally true sentences that authors, readers and critics might *assert* about fictional characters features in fictional works, akin to statements like 'Stephen King wears glasses', and 'Stephen King prefers baseball to soccer'. We can call (4)–(7) *fiction-external* sentences put forth by authors, readers and critics.

The success of irrealism therefore hinges on how compelling, systematic, and general its non-face-value treatment of the various problem sentences is, especially the literally true (4)–(7).<sup>12</sup> The irrealist might analyze these in turn as follows:<sup>13</sup>

- (4') There exists a body of fiction according to which Harry Potter is a specific character.
- (5') J.K. Rowling wrote a body of fiction according to which Harry Potter is a specific character.
- (6') For some property *P*, Stephen King entertains a propositional attitude with the content that Harry Potter has *P*.<sup>14</sup>
- (7') More people think about Harry Potter than about Sparrowhawk.

We can analyze 'think about' further based on (6'):

- (7'') There are more people who entertain propositional attitudes with the content that Harry Potter has *P* for some property *P*, than there are people who entertain propositional attitudes with the content that Sparrowhawk has *Q* for some property *Q*.<sup>15</sup>

<sup>12</sup> van Inwagen (for example, 1977, 2000) argues that there is no systematic way to provide irrealist paraphrases for the problem sentences (4)–(7). For ways that an irrealist might respond, see Brock (2002), Caplan (2004).

<sup>13</sup> See Sainsbury (2010, 115–151); for some of these examples, his preferred analysis is different than the one listed here.

<sup>14</sup> I follow Sainsbury in assuming that an irrealist can readily account for a name like 'Harry Potter' occurring in a propositional attitude context.

<sup>15</sup> The properties *P* and *Q* may vary from one person to the next.

So far, we have considered four proposals about the nature of fictional characters: irrealism, and three realist positions—Meinongianism, nonactualism and artifactualism. Of these, the nonactualist position relies on a notion of possibility and actuality, which are usually cashed out by reference to possible worlds and the actual world; to understand this position better, it is crucial to enumerate briefly the various stances one might adopt with respect to the nature of possible worlds. According to the nonactualist, Harry Potter is a merely possible object who has a thin face, round glasses, knobbly knees, etc.; so (1) is analyzed as:

(2<sup>''</sup>) There is a nonactual possible world in which Harry Potter has a thin face, knobbly knees, black hair, bright green eyes, wears round glasses held together with Scotch tape.

Sainsbury points out that nonactualism incurs a commitment about the metaphysics of possible worlds. “Nonactualists wish to locate [fictional] objects in possible worlds; so they need to be realists about possible worlds” (Sainsbury 2010, 74). There are two major classes of realist views on offer about the metaphysics of possible worlds:

- extreme realism about possible worlds, *pw-realism* for short, has it that the actual world is one among a plurality of possible worlds that are causally and spatiotemporally isolated from one another.<sup>16</sup> This view takes (2<sup>''</sup>) at face value.<sup>17</sup>
- ersatz realism about possible worlds, *pw-ersatzism* for short, has it that a possible world is abstract, for example, a maximally consistent set of propositions representing a way the world could be;<sup>18</sup> this yields the following analysis of (2):

(2<sup>'''</sup>) There is something abstract, a set of (maximally consistent) propositions representing Harry Potter as having a thin face, knobbly knees, black hair, bright green eyes, wearing round glasses held together with Scotch tape.

<sup>16</sup> For a long time, Lewis (1973, 1986) remained the lone proponent of *pw-realism*.

<sup>17</sup> There is a problem, however (Sainsbury 2010, 85–87): in the novels, Harry Potter’s isn’t given a complete description, down to his last detail about sock color; Potter is thus incomplete. But all possible objects are complete (they have the same ontological status we do, it’s just that some of them are nonactual, inhabiting merely possible worlds). One of the more tenable options for the nonactualist is to relate the incomplete Harry Potter to various possible objects—Potter-surrogates—that have all the properties that the novels ascribe to Potter, but are complete (down to the last detail about sock color). Given that on this option, the nonactualist has to quantify over Potter-surrogates to account for (2), she cannot take (2<sup>''</sup>) at face value in the end.

<sup>18</sup> *Pw-ersatzism* has had many proponents. Adams (1974) held this particular, proposition-based version of *pw-ersatzism*.

Sainsbury (2010, 222, fn. 8) points out that pw-ersatzism collapses into a view according to which fictional characters are abstract objects (in our case: a representation comprising propositions about Harry Potter, for example). “For then the true metaphysical nature of a supposedly nonactual fictional character is actual”: an actual set of propositions. It is only in conjunction with pw-realism that nonactualism offers a distinct alternative to a theory like artifactualism, according to which Harry Potter is abstract.

Nonactualism therefore comes in a package bundled with pw-realism, a controversial proposal. In addition, we are considering two other realist contenders—Meinongianism and artifactualism. Analyzing negative existential claims like (1) poses a challenge to all of these views. Accounting for (2)–(8) poses various degrees of difficulty to the realist contenders. Let us now turn to how much weight (5) carries, and more generally, how important it is to maintain, as artifactualism does, that authors create fictional characters.

## 2. THE IMPORTANCE OF AUTHORIAL CREATION

Why insist that authors *create* fictional characters? It does seem natural to say (5):

(5) Harry Potter was created by J. K. Rowling.

Artifactualism, positing Harry Potter as an artifact created by J. K. Rowling, takes (5) at face value. But we have already seen that this is not our only option; we could also accept an irrealist analysis of (5) that doesn’t take it at face value: “J. K. Rowling wrote a body of fiction in which Harry Potter is a specific character”. Quite independently of irrealism, several philosophers have had serious qualms about taking (5) at face value: Brock (2010, 338) sets out to “explain why creationism about fictional characters [the view that fictional characters exist by being created by their author(s)] is an abject failure. It suffers from the same problem as theological creationism: the purported explanation is more mysterious than the data it seeks to explain” because it cannot offer a satisfactory account of the spatial and temporal dimensions of fictional characters, for example, their moment of creation. Yagisawa (2001, 154) argues that the most influential creationist views (by Searle and van Inwagen) “are ultimately unsuccessful in establishing creationism”; more generally, he thinks no view on which fictional characters exist can do justice to our intuition that a claim like “Harry Potter doesn’t exist” is true and is entailed by the true “Harry Potter is a fictional character”. In the light of such doubts about creationism, it is worth homing in on an argument for artifactualism (a form of creationism), showing that it is the best



form of realism one could adopt *precisely because it incorporates authorial creation*.<sup>19</sup> The goal of this section is to expound such an argument.

How might the various forms of realism handle (5)? On this point, artifactualism shows a clear edge relative to its two rivals. A negligible point of advantage is that according to neither rival theories is Potter *created*—going from nonexistent to existent. According to the Meinongian, Potter isn't created—brought into existence—because he doesn't exist (he just is). And according to the nonactualist, Potter had existed all along as a merely possible object and continues to exist as a merely possible object after the novels are written. According to Sainsbury (2010, 61–63, 82–85), the real advantage of artifactualism concerns its response to the so-called *selection problem*: upon introducing the name 'Harry Potter' in her novel, how does J.K. Rowling manage to select one rather than another among the countless candidate objects? According to Meinongianism, there are countless nonexistent candidates; according to nonactualism, there are countless merely possible, nonactual candidates. Sainsbury (2010, 63) doesn't see “how a Meinongian can offer any sensible account of how an author's or reader's thoughts are supposed to engage with one rather than another nonexistent entity”. We are about to see that a more decisive objection emerges against the Meinongian once we consider the difficulties that the *nonactualist* encounters when it comes to the selection problem and other problems.

In the “Addenda” to his “Naming and Necessity” lectures, Kripke (1972, 156–7) motivates two theses for expressions like ‘unicorn’ and ‘Harry Potter’:

- The *metaphysical thesis*: There is no basis for counting any *merely possible* object as Harry Potter, Sherlock Holmes, a unicorn, etc.
- The *epistemological thesis*: There is no basis for counting any *actual* object as Harry Potter, Sherlock Holmes, a unicorn, etc.

In the metaphysical thesis, Kripke's target seems to be the nonactualist. At the end of this section, we will see, however, that both theses bear on Meinongianism also. Along the way, we will also see that the two arguments are at root intimately connected.

Behind the metaphysical thesis is what we might call the *insufficient-specificity problem*.<sup>20</sup> The Harry Potter novels specify many details about Harry; but they

<sup>19</sup> An argument for fictional characters as objects created by people is noteworthy in the light of Brock's (2010, 340–342) criticism. He calls this the Fundamental Thesis: “Fictional characters, to the extent that there are any, are genuinely created by the authors of the works in which their names (or designating descriptions) first appear.” Brock then remarks that “arguments in support of the fundamental thesis are almost completely lacking”. In this Section, I set out to produce precisely this sort of argument.

<sup>20</sup> Kaplan also emphasizes insufficient specificity as an obstacle to naming nonexistents (1973, 506; 1989, 609).

also leave a lot of other details unspecified, for example, which of various parental cells Harry came from. Due to such lack of specificity in the novels, we have no basis for deciding between two distinct merely possible candidates (they originate from distinct sperms, say) that are just like Harry is described in the novels, which of them is Harry Potter. Notice that it is in part due to insufficient specificity in the novels that Sainsbury's selection problem arises—for the Meinongian as well as the nonactualist.

The epistemological thesis turns out to generate an even deeper problem for the nonactualist, one that we shall see (at the end of this section) affects the Meinongian also. Behind the epistemological thesis is what we might call the *coincidental-resemblance problem*, which Kripke discusses in connection with the mythical species of unicorn:

...the mere discovery of animals with the properties attributed to unicorns in the myth would be no means to show that these were the animals the myth was about: perhaps the myth was spun out of whole cloth and the fact that animals with the same appearance actually existed was mere coincidence. In that case, we cannot say that the unicorns of the myth really existed; we must also establish a historical connection that shows that the myth is *about* these animals. (Kripke 1972, 157, emphasis in the original)

Kripke is making two points here: even if we find animals qualitatively like the unicorns of the myth, that wouldn't justify counting them as unicorns given (i) the lack of historical connection between the newly found species and the use of the expression 'unicorn'; and given that (ii) the unicorn myth was "spun out of whole cloth", not created in the right way, to make the term apply to the newly found species. The upshot of (i) and (ii): we would have no more than mere qualitative coincidence between unicorns as described in the myth and the actual species discovered. And for a proper name, reference takes more than coincidental resemblance, so we don't have any candidate actual objects to count as unicorns.<sup>21</sup>

In the case of the expression 'unicorn', the coincidental-resemblance problem thus arises as a result of two distinct problems: (i) *historical unconnectedness* and (ii) *unsuited mode of introduction*. Pure myth-making mode and pure fiction-writing mode both give rise to expressions that aren't introduced in the right way to refer to actual objects. It is natural to expect the *intentions and beliefs* of language users to be highly relevant in determining the mode in which they introduce expressions of their language. On this point, it is customary to note a key difference between myth and fiction (which I will take for granted for the purposes of this paper):

<sup>21</sup> Kaplan quotes Harry Deutsch: "reference is no coincidence" (Kaplan 1989: 608).

The difference between authors and myth-makers is one of propositional attitude: authors *make-believe* their works of fiction, whereas myth-makers do not make-believe their myths; rather, they genuinely *believe* their myths. (Caplan 2004, 334, emphasis in the original)<sup>22</sup>

According to this, those who created the myth of the unicorn had unicorn-related *beliefs* (not just pretended beliefs). Still Kripke does raise the issue of an unsuited mode of introduction: myth-spinning. Right after the passage above, Kripke (1972, 157–158) repeats the same point with respect to ‘Sherlock Holmes’ also: “it is theoretically possible though in practice fantastically unlikely, that Doyle was writing pure fiction with only coincidental resemblance to [an] actual man”. A crucial consideration emerges from these fleeting remarks about unicorns and Sherlock Holmes: given (ii) the way the myth/fiction was created, and (i) the fact that we encounter historical unconnectedness, the result is that we find no more than coincidental resemblance to actual objects.

The unsuited-mode problem would arise even if we had at hand a myth or a novel specifying mythical beings/characters completely, down to the last bit of information about sock color and origin (it would be mind-numbing to read such a novel).<sup>23</sup> So even in special cases of names from complete fictions in which the metaphysical thesis is circumvented, the epistemological thesis would still present problems. (Given the focus of this paper, in what follows, I will focus on characters from fiction, setting myths and mythical beings to the side; the points I make about the various problems can be generalized to names from myths also.)

Both theses and all the problems considered so far have taken it for granted that the candidate objects to count as Harry Potter are *concrete, spatiotemporal objects*. It is therefore well to keep this qualification in mind. For example, for (ii) we get: the fiction-writing mode in which the expression ‘Harry Potter’ had been introduced into the language is unsuited for the name to refer to an actual concrete, spatiotemporal object. For (i) we get: actual, concrete, spatiotemporal objects as potential referents for the name are historically unconnected to the introduction and subsequent use of ‘Harry Potter’.

It’s crucial to note that of the two problems (i) and (ii), unsuited mode of introduction is the more fundamental one, explaining historical unconnectedness *of the relevant sort*: for all we know, there could be an actual person who provided inspiration for Rowling’s Harry Potter; this would make for *a* historical connection between Rowling’s use of the name and the actual boy. But it wouldn’t be the *reference-determining* kind of historical connection we’re interested in, the sort of historical connection that would circumvent the coincidental-resemblance

<sup>22</sup> See also Salmon 1998; Braun 2005.

<sup>23</sup> Kaplan (1989, 609) makes this point.

problem by determining the reference of ‘Harry Potter’. Rowling’s authorial intentions, her fiction-writing mode precludes historical connections that are reference-fixing.<sup>24</sup> Kripke (1972, 92) also gives an example of an irrelevant kind of causal/historical connection: the “causal chain from our use of the term ‘Santa Claus’ to a certain historical saint”; despite such a link, children, when they use the name ‘Santa Claus’ “by this time probably do not refer to that saint”.

Being an actual person who is the spitting image of Harry Potter as he is described in the Rowling novels would make for no more than coincidental resemblance to Harry Potter. Here is why: J. K. Rowling’s intention was to introduce the name ‘Harry Potter’ for a fictional character rather than an actual person who fits a certain set of descriptions. And for a name to refer to an actual person takes more than coincidental resemblance; reference is shaped (i) in part by causal-historical connections between uses of the name and an object (whether that be a concrete or an abstract object), and (ii) in part by the mode of introduction. Given that (ii) Rowling’s intention was to create a *fictional character* rather than refer to a flesh-and-blood person with introducing ‘Harry Potter’, (i) the name ‘Harry Potter’ was never historically linked (in the relevant way) to an actual orphaned boy wearing glasses, with a Z-shaped scar on his forehead, growing up in suburban England learning wizardry in a boarding school, and so on, and the name cannot refer to any actual concrete boy with spatiotemporal dimensions.

Not only is the unsuited-mode problem more fundamental than (i); it is also more general. Notice that it readily generalizes to concrete, spatiotemporal objects of all sorts, merely possible ones included; this way, we get:

*the unsuited-mode problem generalized:* the fiction-writing mode of introducing proper names into the language is unsuited for them to have as their reference concrete, spatiotemporal objects, whether they be actual or merely possible.

It is well to generalize in the same way the coincidental-resemblance problem also:

*The coincidental-resemblance problem generalized:* there is no more than mere qualitative coincidence between concrete, spatiotemporal objects (whether they be actual or merely possible) and fictional characters as described in works of fiction.

<sup>24</sup> Notice that I am assuming here that authors have *the final word* on whether they are introducing a name for a historical figure or a fictional character. Suppose that Tolstoy, upon asked about Napoleon in his *War and Peace* were to have sincerely said: “I intended the figure of Napoleon in *War and Peace* to be a fictional character; I drew a great deal of inspiration from the French military leader, but still, resemblance between the character and the historical figure is pure coincidence”. In this imagined scenario, the position I am assuming is that Tolstoy would have introduced ‘Napoleon’ in *War and Peace* in fiction-writing mode, so the reference of the name would not have been the historical figure, and historical connections to the first Emperor of the French would not have been of the relevant, reference-fixing sort.

Therefore, as we dig deeper, the pair of problems behind the epistemological thesis turn out to target nonactualism.

As before, in the case of 'Harry Potter', the unsuited-mode problem generalized underlies the generalized coincidental-resemblance problem. And both apply to the metaphysical thesis also: the generalized unsuited-mode problem provides the following *additional* reason for holding the metaphysical thesis. *If the character of Harry Potter is not fully specified in the novels, then what grounds do we have at all for choosing between two distinct merely possible concrete, spatiotemporal objects which to count as Harry Potter when, given J. K. Rowling's fiction-writing mode of introducing 'Harry Potter', it would be a matter of sheer coincidental resemblance for the name to refer to either of those candidate objects? With respect to names from fiction, the unsuited-mode problem (and in its wake, the coincidental resemblance problem) therefore raises a key issue underlying both the metaphysical and the epistemological theses discussed by Kripke; this is a striking detail to bring to the surface given that Kripke mentions the unsuited-mode problem in passing only (saying no more than the two half-sentences quoted above), devoting far more attention to the metaphysical thesis.*

Just how bizarre the idea of reference based on coincidental resemblance is—the conception of reference for 'Harry Potter' to which the nonactualist is committed—can be brought out based on considerations about *nonfictional* names that fail to refer. The French astronomer Le Verrier put forth a hypothesis about the existence of an intra-Mercurial planet which he named 'Vulcan', to explain perturbations in the orbit of Mercury. There were various independent sightings mistakenly believed to be of Vulcan before enthusiasm dwindled; By 1916, Einstein's general theory of relativity confirmed that the perturbations were produced by the gravitational field of the Sun; there was no intra-Mercurial planet at all; the Vulcan-hypothesis was refuted; 'Vulcan' turned out not to refer to anything.

What about a counterfactual situation in which the Vulcan-hypothesis is a success story? Imagine a counterfactual scenario with the laws of physics slightly different, and there being an intra-Mercurial planet affecting the orbit of Mercury; Le Verrier puts forth his hypothesis; there are sightings converging on the planet, which comes to be called 'Vulcan', the name featured in Le Verrier's prior hypothesis. But that is not *our* term 'Vulcan' that comes to name the counterfactual planet, but a different one. It is preposterous to think that in coining the name in the actual world, Le Verrier managed to name *that counterfactual object even though his naming attempt failed in the actual world*. 'Vulcan' might have been a success story just as 'London' might have been introduced as a name for a river instead of a city; but all that is irrelevant to how and whether these strings, as parts of *our* language, were introduced and subsequently used.<sup>25</sup> Le Verrier

<sup>25</sup> See Kripke (1971, 145; 1972, 77, 102–3, especially fn. 51).

strove to name an actual concrete, spatiotemporal object; due to his failure to do so, he didn't *by coincidence* name a nonactual concrete, spatiotemporal object (as the nonactualist would have it); doing so was no part of his intention. So 'Vulcan' doesn't refer to any concrete objects in any counterfactual situations. Kaplan (1973, 506–508) makes this point eloquently with respect to a fictional name like 'Pegasus'. But what is far more interesting is that the point holds for 'Vulcan'! We can say the following about this name, as well as other proper names intended for concrete objects or for fictional characters:<sup>26</sup> *if it cannot make it here, it cannot make it anywhere*. If the name doesn't refer to a concrete, spatiotemporal object here, in the actual world, it doesn't refer to such an object in other possible worlds either. Elsewhere (Zvolenszky 2007), I call this the *inverse-Sinatra principle for proper names*.<sup>27</sup>

The inverse Sinatra principle is quite general, covering names like 'Vulcan', 'Pegasus', and 'Harry Potter'. And the reason why these names cannot make it anywhere given that they cannot make it here (in the actual world), is because nonactual concrete objects are, at best, coincidentally similar to the descriptions given for Vulcan, Pegasus and Harry Potter. We thus have a nonfictional variant of the coincidental resemblance problem: there is no more than mere qualitative coincidence between merely possible concrete, spatiotemporal objects and nonreferring names featured in failed hypotheses.

Notice that 'Vulcan' and 'Harry Potter' differ in one crucial detail: for the case of 'Vulcan', the unsuited-mode objection doesn't arise. Le Verrier's intention had been to introduce 'Vulcan' for a concrete, spatiotemporal object; so a historical connection, if there had been one, linking uses of the name to an actual concrete object, could have served to fix the reference of 'Vulcan', circumventing coincidental-resemblance-related qualms. A historical connection can be secured in the actual world only—there is absolutely no historical connection between *our* use of 'Vulcan' and a merely possible concrete, spatiotemporal object. And in the absence of an actual historical connection, qualms about coincidental resemblance do arise, leading to the metaphysical thesis about 'Vulcan': if the specification of Vulcan isn't complete, allowing that several distinct merely possible concrete objects fit the specification equally, then we have no basis for counting any one of them as Vulcan. (Notice that here, as before, my argument leading to the metaphysical thesis for Vulcan was crucially linked to considerations about coincidental resemblance and historical unconnectedness,

<sup>26</sup> Even an irrealist about fictional characters can, based on the considerations about Vulcan and unicorns above, accept the inverse-Sinatra principle.

<sup>27</sup> Frank Sinatra sang about New York City: "If I can make it there, I'll make it anywhere".

which were originally identified behind the other thesis—the epistemological one. With respect to ‘Vulcan’, too, we see that the two theses are intimately connected.)

The foregoing observation allows us to highlight a more general point of advantage for the artifactualist position over both Meinongianism and nonactualism.

According to artifactualism, Harry Potter is an actual object; it’s just that unlike concrete objects like the Big Ben, Harry Potter is abstract. Yet the fact that he is an actual artifact makes room for a certain kind of causal-historical dependence on the physical world: in the 1990s, J. K. Rowling’s creative activities bring it about that Potter is an actual (rather than a merely possible) abstract object. The sort of dependence in place allows Harry Potter qua abstract artifact to be the kind of referent for Rowling’s name ‘Harry Potter’ with respect to which issues having to do with historical unconnectedness and, in turn, coincidental resemblance, and, in turn, the epistemological thesis, do not arise. (Notice that before, we noted that for names of fictional characters, no historical connection to concrete, spatiotemporal objects is of the relevant, reference-fixing sort. Meanwhile, the point made here is that for the artifactualist, a historical connection to an *abstract artifact* is precisely what fixes the reference of ‘Harry Potter’.)

By contrast, alternative realist accounts that make Harry Potter a concrete object whose existence does not causally depend on us either because the object is nonexistent (according to Meinongianism) or because it is nonactual (according to nonactualism), face a challenge. First, these theorists have to explain why those objects are candidates of the right ontological status to count as the referents of ‘Harry Potter’. As we have already seen, on this point, the nonactualist founders already. The Meinongian can get past this hurdle: he may suggest that his nonexistents are objects of thought and hence have just the right sort of ontological status to be suitable targets of authors’ intended reference. But on the next hurdle the Meinongian stumbles: if his nonexistent objects are of a suitable sort as objects of fiction-writing, what historical connection is there to account for Rowling’s ‘Harry Potter’ referring to one of countless nonexistent candidate objects (each equally faithful to the way Potter is depicted in the novels but varying in details left unspecified—about sock color, etc.)? The Meinongian cannot provide such a historical connection: causal-historical connection between his timelessly nonexistent objects and actual concreta (like authors) is extremely problematic, downright unintelligible even. And because of historical unconnectedness, the Meinongian is confronted with qualms about having to work with no more than coincidental resemblance between Harry Potter as specified in the novels, and various qualitatively identical Meinongian nonexistents. And, on the one hand, coincidental resemblance does not suffice for reference, according to the epistemological thesis; and, on the other hand, with insufficiently



specified characters like Harry Potter, coincidental resemblance leaves room for the metaphysical thesis (and also the selection problem) to arise.<sup>28</sup>

Once fleshed out, Kripke's (1972) fleeting remarks about fictional characters can be summarized as follows: qualitative resemblance is insufficient to determine the reference of a proper name; a causal-historical connection between names and their referents is necessary to determine to whom or to what proper names refer. For names of actual objects like 'J.K. Rowling' and 'London', this overarching lesson transparently emerges from the second lecture of *Naming and Necessity*. It is considerably less transparent that Kripke reiterates the very same lesson for names of fictional characters. Of the forms of realism considered, artifactualism is the only one that can heed this lesson.

### 3. DEFLECTING THE CATEGORY-MISTAKE OBJECTION

Sainsbury, an advocate of irrealism, agrees with the conclusion of the previous chapter: among realist contenders, artifactualism has the edge. Unlike Yagisawa (2001), he doesn't think that the major challenge artifactualists face concerns accounting for the truth of negative existential claims like "Harry Potter doesn't exist". Instead, Sainsbury (2010, 111) writes:

I see the problems for abstract artifact theory lying elsewhere... They have the form: on abstract artifact theories, fictional characters just are not the kinds of things we want them to be. We want them to be as they are said to be in the stories, to be detectives and to play the violin, but they are said to be something of an entirely different kind.

...

Fictional characters do not have the properties they are ascribed during their creation. This is mysterious: Conan Doyle stipulates that Holmes wears a deerstalker, there is such an entity as Holmes, yet that entity does not end up having (i.e. exemplifying) the property of wearing a deerstalker. He does end up having (exemplifying) a genuine property, that of encoding wearing a deerstalker, but this is not

<sup>28</sup> This line of argument brings to the fore why the only abstract-theory contender we considered for fictional characters was artifactualism: it is the only view according to which Harry Potter is created and hence historically linked to goings on in the actual world. Platonism, a theory according to which Harry Potter is a *timelessly existing* abstract object (akin to numbers, sets), would, like Meinongianism and nonactualism, run into problems with historical unconnectedness and hence coincidental resemblance, and, in their wake, the metaphysical and epistemological theses. For an attempt to combine the advantages of artifactualism and Meinongianism, see Zalta's (2000, 2006).



a property that's intellectually accessible to most authors. People can, of course, fail to understand what they are doing, but it's surprising to be told that so many authors, perhaps all, fail so often and so seriously.

Sainsbury is here relying on a distinction between exemplifying and encoding originally suggested by Meinong's student Mally (1912): a concrete object like J.K. Rowling doesn't encode any properties; but she does *exemplify* being British and fails to exemplify wearing glasses. Meanwhile, Harry Potter *encodes* wearing glasses and being British, but exemplifies neither of these properties. He does, however, exemplify being abstract and being a fictional character.

Part of Sainsbury's objection then is that according to artifactualism, fictional characters are of the wrong ontological category—abstract rather than concrete—to exemplify the sorts of properties ascribed to them by the authors who create them. I call this *the category-mistake objection*. A consequence of the category-mistake objection is that artifactualism attributes massive error to those who create, read and discuss fictional characters.

My aim is to show that the strategy behind the category-mistake objection, if it were to work, would show far too much with respect to a broad range of metaphysical debates. The strategy is therefore objectionable. I will formulate three arguments to show that the category-mistake strategy does not withstand scrutiny.

My first argument is about the metaphysics of possible worlds. In Section 1, we distinguished two positions in the debate about the nature of possible worlds: realism versus ersatzism about possible worlds (pw-realism and pw-ersatzism). According to pw-ersatzism, possible worlds are abstract, for example, maximally consistent sets of propositions representing ways the world could be. And merely possible individuals are likewise abstract (as Sainsbury acknowledges): representations comprising propositions about the individual. Now, when I consider a counterfactual scenario in which I dye my hair green today, I am ascribing to myself the property of having green hair, or so it seems to me when I reflect on my mental episode. Yet the category mistake objection could be raised here: according to pw-ersatzism, possible objects are of the wrong ontological category—abstract rather than concrete—to exemplify the sorts of properties ascribed to them by those who entertain counterfactual scenarios. This objection would apply to all forms of pw-ersatzism, regardless of whether they construe worlds in terms of states of affairs, universals or sentences. On all these versions, possible objects are the wrong kinds of things to be ascribed the properties we ordinarily ascribe to them. Anyone who thinks pw-ersatzism cannot be dismissed quite so

quickly has reason to consider the strategy behind the category-mistake objection (as targeting artifactualism as well as pw-ersatzism) specious.<sup>29</sup>

Another point casts further doubt on the category-mistake strategy. The category-mistake objection against pw-ersatzism, if it were to work, would seem to leave the rather controversial position of pw-realism, famous for eliciting incredulous stares (Lewis 1973, 86), as the only realist account of possible worlds. According to pw-realism, the actual world is one among a plurality of possible worlds that are causally and spatiotemporally isolated from one another. A moment's further thought reveals that an objection closely related to the category-mistake objection affects pw-realism also. If an ordinary speaker were asked if she thought there are countless merely possible worlds and countless merely possible objects, and if she thought such things have the same ontological status as the actual world and actual objects, respectively, she would answer in the negative to both questions. Hence the incredulous stare that confronts pw-realism. Yet, contrary to people's intuitions, pw-realism posits that possible worlds have the same ontological status as that enjoyed by the actual world, and possible concrete objects have the same ontological status enjoyed by actual concrete objects. Call this the *mistaken-ontological-status objection to pw-realism*. It is unclear why this objection should have any less force than the category-mistake objection against pw-ersatzism. But if that objection were effective against pw-ersatzism while the mistaken-ontological-status objection were effective against pw-realism, then irrealist accounts of possible worlds would be left as the only alternatives standing. This conclusion seems much too quickly and easily obtained for irrealists about possible worlds (including Sainsbury). The pair of objections seem, from the outset, to rig the stakes against *all* forms of realism about possible worlds. Anyone who thinks that realism about possible worlds cannot be dismissed quite so easily has reason to consider at least one of the two objections specious. Until the irrealist about possible worlds provides special reasons that discredit the mistaken-ontological-category objection, both objections remain suspect.

My second argument is intended to show that for someone who finds a form of pw-ersatzism (a not unpopular view about the metaphysics of possible worlds) independently plausible, there is little reason to *resist* admitting fictional characters as abstract objects, the category-mistake objection notwithstanding. Here is

<sup>29</sup> An argument from authority (whatever its merits might be): interestingly, while Lewis (1986) carefully considered a long list of arguments against pw-ersatzism, he did not address the category-mistake objection against it.

Of course, one person's *modus tollens* is another *modus ponens*; I have motivated the following conditional: if the category-mistake objection is effective against artifactualism, then it is effective against pw-ersatzism. I have taken the *modus tollens* direction and concluded that the objection is ineffective against artifactualism. Someone else might take the *modus ponens* direction and conclude that the objection is a new and effective one against pw-ersatzism.

why. Imagine a certain spool of yarn I knit into a sock: Sock1, based on a specific set of knitting instructions. Imagine another specific, actual spool of the same yarn that I could have used to knit a qualitatively identical (or very similar) pair to Sock1: Sock2. As things stand, I never got around to knitting Sock2. So Sock2 doesn't exist, but it might have.<sup>30</sup> According to the pw-ersatzist, Sock2 is abstract: we might take it to correspond to a set of propositions representing Sock2 as having a certain color, shape, size, pattern, person knitting it; the proposition set can then be said to encode color, shape, knitter, etc., and exemplify being abstract and consisting of propositions. This set of propositions actually exists; but the scenario it represents is unactualized (because I never get around to knitting Sock2). Now consider the fictional sock that is featured at the end of the second Harry Potter novel: Harry pulls off one of his "slimy, filthy" socks, tricking Lucius Malfoy into unwittingly giving it to his long-suffering house elf, Dobby, thereby releasing Dobby from serving the Malfoy family (house elves are freed when their masters give them an article of clothing). J.K. Rowling doesn't give this sock a name, but for easy reference, let's call it DobbySock. According to artifactualism, J.K. Rowling created DobbySock; she specified it as having been worn by Harry Potter, as being slimy and filthy, but she didn't say what color it was. Whatever form of pw-ersatzism we might opt for, we can go the same way with DobbySock; for example, we can take DobbySock to correspond to a set of propositions representing it as being filthy, slimy, etc.; this set of propositions encodes DobbySock as being filthy, slimy, etc. (in the second Harry Potter novel); and the same set exemplifies being Rowling's creation, a fictional character, a famous fictional character even. We must realize that by taking these parallel approaches to Sock2 and DobbySock, there isn't that great a difference in the nature of the merely possible Sock2 and the artifact DobbySock: both correspond to sets of propositions representing socks, encoding certain properties (like being a sock) and exemplifying others (like being abstract, containing propositions). One noteworthy difference is that the first set *encodes* being created by me (the knitter), while the second set *exemplifies* being created by J.K. Rowling. But this is a difference we expect, and the similarities are otherwise striking. For a pw-ersatzist, with merely possible objects on board, it would be *ad hoc* to resist what is mostly parallel treatment for fictional characters: a form of artifactualism. Given how costly such an *ad hoc* move would be, the pw-ersatzist should embrace artifactualism and not worry about the category-mistake objection.

It is well to address three worries at this point. First, there is a crucial difference between Sock2 and DobbySock according to someone who combines

<sup>30</sup> Notice that 'Sock2' is a special name to which the inverse Sinatra principle discussed in Section 2 does not apply. And that's all well: 'Sock2' doesn't make it here but makes it in the possible worlds in which it gets knit.

pw-ersatzism and artifactualism: (on at least one plausible view of propositions), the proposition set for Sock2 exists timelessly, as do the pw-ersatzist's possible worlds; by contrast, DobbySock is an artifact and hence not a timeless existent. Why should the pw-ersatzist be moved to admit the latter kind of beast then, an object that is unlike her possible worlds and objects? Three reasons: (i) beyond this difference, there are crucial similarities between Sock2 and DobbySock, ones that make it plausible to treat both as abstract; (ii) operas, novels, the institution of marriage, etc. are overwhelmingly plausible candidates for abstract artifacts already; so the burden of providing a workable alternative is on those who want to deny that these are abstract artifacts; (iii) with operas and novels on board as abstract artifacts, between the Platonist and artifactualist alternatives, the latter is a far more tenable choice. For the Platonist view (mentioned in footnote 3 above)—according to which DobbySock is a timelessly existing abstract object—is affected by the selection problem, insufficient specificity, unsuited mode of introduction and coincidental resemblance; problems (discussed in Section 2) that artifactualism (singularly among realist contenders) avoids.

Second, notice that in contrasting encoding and exemplifying above, I have talked about *proposition sets* encoding and exemplifying properties like being knit by me and created by J.K. Rowling. Proposition sets represent ways the world might be; they are representational devices. And “[t]he distinction between encoding and exemplifying is one that is properly available for representational vehicles, but that’s not what fictional characters are. They are what’s represented”, Sainsbury (2010, 112) objects. The worry is that fictional characters qua abstract objects aren’t the right sorts of things to be representational devices and to be encoding properties. This worry is easily responded to: although possible worlds seem at first like really big particular objects, like all-encompassing, gigantic galaxies, the ersatz-realist does not balk at construing them as sets of propositions or as structural universals<sup>31</sup>. In the same way, the artifactualist should not worry about taking fictional characters to be sets of propositions (or as structural universals). That it is the proposition set about Harry Potter that encodes and exemplifies various properties is not a problem given that fictional characters correspond to such proposition sets. On the version of pw-ersatzism we are considering, “Sock2 encodes shape, size, pattern and exemplifies being abstract” is loose talk for “The ‘Sock2’ proposition set represents Sock2, encoding shape, size, pattern, and exemplifying being abstract”. Likewise, on the version of artifactualism we are considering, “DobbySock encodes being filthy and slimy and exemplifies being abstract and created by J.K. Rowling” is loose talk for “The ‘DobbySock’ proposition set represents DobbySock, encoding filthiness,

<sup>31</sup> A paradigmatic example of a structural universal is *being a water molecule*: for an object to instantiate this universal, it has to have the right kinds of parts in the right kind of arrangement.

sliminess, and exemplifying being abstract and created by J.K. Rowling.” Perhaps some of the pw-ersatzist’s alternatives are conceptually more satisfying in some way than proposition sets; but the point stands: whatever kind of abstract objects the pw-ersatzist might posit as her possible objects, she has good reason to extend her theory to some very similar beasts: fictional characters as abstract artifacts. And her choice of construal for these objects can then accommodate the encoding/exemplifying distinction in much the same way as the proposition set construal did.

A third worry arises: aren’t we multiplying abstract objects that are qualitatively identical to one another? Imagine a merely possible sock that is qualitatively identical to DobbySock, as specified in the second Harry Potter novel, call it JustLikeDS. The ‘JustLikeDS’ proposition set encodes the same properties as the ‘DobbySock’ proposition set encodes. The two sets exemplify some of the same properties: being abstract, being sets, consisting of propositions. Now, isn’t it an extravagant proliferation of objects to hold that with JustLikeDS already in existence, J.K. Rowling creates a qualitative duplicate, DobbySock, upon conjuring up the second Harry Potter novel? We can see that this outcome is not worrisome at all if we reflect on some perfectly ordinary scenarios that are analogous.

Consider another abstract object, say, Mozart’s opera *The Magic Flute*. Consider the collection of musical chord sequences, timing, order, etc. for the various singers and instruments, which the total score of *The Magic Flute* comprises. This is a type, which is a paradigmatic instance of an abstract object that can have specific performances of the opera as its tokens. Now, *The Magic Flute* qua abstract type was created by Mozart. But now consider a type—call it JustLikeMF—of chord sequences, instructions, various specifications qualitatively just like the score of *The Magic Flute*. If we take sets, properties and propositions to be timelessly existing abstract objects (a widely held position), then clearly, a type such as JustLikeMF should be regarded as an abstract object, plausibly, a timelessly existing one. But then when Mozart came along, he ended up proliferating qualitatively identical types by writing the score of *The Magic Flute*, in addition to the timelessly existing JustLikeMF. Proliferation of this sort is inevitable, yet it isn’t taken as cause for concern for those who posit types as abstract, and musical pieces as types created by composers.

Consider a similar example: the swiftly created and enacted new Hungarian constitution (“Fundamental Law” it’s called) didn’t always exist; it came into existence in 2011 only, when it was drafted; indeed, beforehand, many considered it unfathomable that an object like the Fundamental Law should ever be created; but it was. The Fundamental Law is a type that can have instances: printed and electronic copies, a reading event of the text. The Fundamental Law didn’t exist before the current government came into power, but it exists now. Yet a qualitatively identical type, an ordered sequence of propositions, is

plausibly an abstract object that existed well before 2011, if not timelessly<sup>32</sup>. Proliferation of this sort is inevitable if we want to maintain that the Fundamental Law is an artifact created in the recent past while types are abstract also. And if proliferation is no cause for concern here, it isn't worrisome in the case of DobbySock and JustLikeDS either.

Consider a third example: words being added to the English vocabulary. For example, a fairly recent addition to the English language is the expression 'cot potato', meaning a very young child who spends a lot of time watching television. The expression type 'cot potato' can have hand-written, typed, electronic, spoken, mouthed or signed tokens. Linguists tend to take for granted that expression types are abstract objects, specifically, abstract artifacts that didn't always exist. But (relative to 'cot potato') a qualitatively identical phonological type, orthographic type, and semantic type qua abstract types have been around for much longer,<sup>33</sup> so with the addition of new words like 'cot potato' to the English language, we get a the very same kind of proliferation that DobbySock and JustLikeDS had presented; and this sort of word proliferation is rampant: for any expression type of any language, there is a qualitative duplicate that is an antecedently existing abstract object. If that isn't worrisome, nor is the case of DobbySock and JustLikeDS.

The upshot of these examples is that proliferation of qualitative duplicates is inevitable for abstract artifacts across the board. If (like many theorists) we still want our ontology to make room for works of art, social and legal institutions, games, words within a language, traditions, festivals, religions as abstract artifacts, then we should have no qualms about including fictional characters on the list. And overall, the upshot of my second argument has been that for *pwersatzists*, resisting artifactualism would be an *ad hoc* move.

My third argument is that the category-mistake objection can be extended to a broad range of objects that are commonly regarded as abstract artifacts. If we don't balk at the objection there, we should pay no head to it with respect to artifactualism either. Were those who were coming up with and modifying the rules of chess thinking of themselves as creating something abstract? Were the writers of the Fundamental Law of Hungary thinking of themselves as creating something abstract? Was Mozart, when composing *The Magic Flute*? Were those who coined the term 'cot potato'? If these people were interviewed, they would

<sup>32</sup> Whether or not we take the ordered sequence of propositions to exist timelessly depends in part on our view of propositions, an issue on which I'd like to maintain neutrality. Either way, the qualitatively identical type enjoys prior existence relative to the Fundamental Law.

<sup>33</sup> I avoid talking about timelessly existing types here for the sake of neutrality on various matters. I want to leave open the possibility that the orthographic type 'cot potato' doesn't exist timelessly because it didn't exist prior to the English orthographic system coming into existence. Likewise, I want to leave open the possibility that the semantic type 'cot potato' doesn't exist timelessly because it didn't exist prior to the existence of television sets.

likely be baffled by the idea that they were aiming at creating abstract objects. Yet regarding the game of chess, the Fundamental Law, *The Magic Flute*, and words of English as abstract artifacts is a plausible option, more plausible than its alternatives, and one that many consider platitudinous. Then why worry about the category-mistake objection against artifactualism?<sup>34</sup>

A defender of the category-mistake objection can make a comeback: all the abstract artifacts considered here are types capable of having concrete, spatio-temporal instances such as specific games of chess, copies of the Fundamental Law, performances of *The Magic Flute*, utterances of ‘cot potato’. The creators of these abstract artifacts *were specifying conditions for the instances of the types they were creating*. But there is a crucial disanalogy between the type-like abstract artifacts just considered and fictional characters qua abstract artifacts: the latter are not the kinds of things capable of having instances. Hence the application of the category-mistake objection to the latter.

The artifactualist can respond to this in three ways. Granted: types can have instances. But it is not so outlandish to think of fictional characters as having something a bit like instances: for example, an opera singer singing Papagena in *The Magic Flute plays* Papagena; though admittedly, she *isn't* Papagena. In the Harry Potter movies, there were several actors *playing* Professor Dumbledore; though admittedly, none of them *was* Dumbledore. We might insist, however that the opera singer and the actors *are*—within the fiction—instances of the characters they play, so the disanalogy isn't as great as the defender of the category-mistake objection makes out.

A second and more substantial line of response from the artifactualist: consider, again, merely possible objects like Sock2, which are plausibly construed within a *pw-ersatzist* framework as, for example, a set of propositions. This proposition set can be instantiated in a sense: it can be *actualized*—it is actualized when I end up knitting Sock2. We can think in terms of this model of actualization for fictional characters like DobbySock qua abstract artifact also. It is just that for reasons explained in Section 2, fictional characters are forever unactualized, indeed, unactualizable. The inverse Sinatra principle sums up a crucial feature of names, fictional names included: if they don't make it here, they don't make it anywhere; that is, if they don't refer to concrete, spatiotemporal objects in the actual world, then they don't do so with respect to other possible worlds either. Fictional characters by their very nature are barred from being actualized. But

<sup>34</sup> I am sympathetic to Thomasson's (1999; 2009, 16) point “that those who accept the existence of such ordinary social and cultural objects as laws, marriages, symphonies, and works of literature themselves are apparently already committed to the existence of created abstracta, so that no special problems arise in accepting created abstracta to account for fictional characters”, and “it is not obviously more parsimonious to do without fictional characters if we must posit abstract artifacts in some other arena, e.g. to make sense of our talk about novels, symphonies, laws of state, and the like.”



this is a feature of theirs due to considerations about coincidental resemblance being insufficient for reference. This feature does not make for a decisive point of disanalogy between them and merely possible individuals. So granted, there is a special reason why in one sense, fictional characters qua abstract objects aren't ever instantiated (that is, actualized); but that need not make them radically different from types.

Third, consider the tradition of the Easter rabbit and its Australian marsupial variant, the Easter bilby. What exactly might be concrete, spatiotemporal instances of these traditions—events involving nest-making, egg-painting, candy-hiding, candy-hunting?—is somewhat puzzling. So fictional characters aren't the only abstract artifacts for which it isn't straightforward what their instances might be. But saying that the tradition of the Easter bunny is therefore not an abstract artifact while the other instantiable types are, invites the challenge for the defender of the category-mistake objection: on what basis will she count among her abstract artifacts the tradition of the Village Halloween Parade in New York City and Mardi Gras festivals, but exclude the tradition of the Easter bilby?

The upshot of the third argument is that the difference between types and fictional characters qua abstract artifacts isn't as great as it initially appears. Yet the category-mistake objection applies to a broad and varied range of types: the game of chess, *The Magic Flute*, words. If the objection is ineffective there, we have little reason to worry about it when it comes to fictional characters conceived as abstract artifacts.<sup>35</sup>

#### 4. CONCLUSION

My aim in this paper has been to motivate taking at face value claims about fictional characters being created by their authors. This requires a form of realism about fictional characters: they are supposed to *exist* once created. Among the contending realist theories about fictional characters that Sainsbury (2010) considered, artifactualism, which takes Harry Potter to be an abstract artifact, is the only one according to which fictional characters are created. Based on Kripke's

<sup>35</sup> Notice that the third argument is not defeated by those who contest the abstract artifact status of one or another among the various candidates I have enumerated. As long as these philosophers are moved to grant abstract artifact status to *some* of the examples mentioned—operas, novels, words or the game of chess—the third argument affects them already. Meanwhile, for philosophers holding that *none* of the items listed in Section 1 as abstract artifacts are in fact such (because they suggest, say, that these aren't abstract, after all), the burden of proof is on them to explain how operas, novels, words, etc. are nonabstract.



(1972) brief considerations about expressions like ‘unicorn’ and ‘Sherlock Holmes’, I constructed a general argument showing that artifactualism is superior to its realist rivals. Sainsbury accepts this conclusion, yet argues that one should ultimately reject realism about fictional characters because artifactualism faces insurmountable difficulties due to the category-mistake objection. I gave three arguments showing that the category-mistake objection is problematic because if it were to work, it would show too much: first, it would show ersatzism about possible worlds to be a nonstarter; second, it would prevent the ersatzist from taking on board fictional characters as abstract artifacts, an *ad hoc* move for her; and third, it would cast doubt on the abstract artifact status of a broad range of social and cultural artifacts like the game of chess, words of English, bodies of law, novels and operas. *Pace* Sainsbury, artifactualism about fictional characters remains unscathed by the category-mistake objection.

In the process of disarming the category-mistake objection, I aimed also to demystify what Harry Potter as an abstract artifact might be. Let me close with two further demystifying considerations. Brock (2010) deems a position like artifactualism “an object failure” because it cannot offer a satisfactory account of the spatial and temporal dimensions of fictional characters, for example, their moment of creation. This line of attack ignores that quintessential examples of abstract artifacts like the institution of marriage and the game of chess are just as difficult to locate in space and time as Harry Potter is. In particular, when each came into existence is a thorny question: there is extensive debate as to which stage of rules for a board game to count as the birth of chess—the 15<sup>th</sup> century, the 17<sup>th</sup>, or the 19<sup>th</sup>? Most likely, no agreement will be reached on this issue; but that hardly casts doubt on the abstract artifact status of the game of chess. Nor has Brock given us reason to doubt the abstract artifact status of Harry Potter.

Consider a related point about criteria for individuating abstract artifacts: does a board game very much like contemporary chess but with different rules about stalemate count as *chess* at an earlier stage? How much meaning change and/or pronunciation change can a *word of English* undergo and still remain the same word? Can the institution of *marriage* be modified such that the two parties may be of the same sex? The considerations that could decide such questions seem unclear and arbitrary. Yet the lack of clear answers doesn’t call into question the abstract artifact status of the game of chess, words of English and the institution of marriage. Abstract artifacts can and do change over time as humans shape and modify them. In the light of this, it is only expected that Harry Potter qua abstract artifact may change as the film adaptations fill in details (for example, the sock Harry sneaks to Dobby is black) or contradict the original novels (for example, in the novel, as Harry gets ready to trick Dobby’s master into giving Dobby a sock, Harry hides the diary of Tom Riddle inside the sock, while in the movie, he hides the sock inside the diary). Or maybe the Harry of the novel is a distinct artifact than the Harry of the movies? Again a clear and nonarbitrary

answer seems doubtful. But those who, upon encountering comparable issues with quintessential abstract artifacts like words of English and the institution of marriage did not speak up, should make peace with Harry Potter as an abstract artifact.<sup>36</sup>

## BIBLIOGRAPHY

- Adams, Robert M. 1974. Theories of Actuality. *Noûs* 8: 211–231.
- Braun, David 2005. Empty Names, Fictional Names, Mythical Names. *Noûs* 39: 596–631.
- Brock, Stuart 2002. Fictionalism about Fictional Characters. *Noûs* 36: 1–21.
- Brock, Stuart 2010. The Creationist Fiction: the Case against Creationism about Fictional Characters. *Philosophical Review* 119: 337–364.
- Caplan, Ben 2004. Creatures of Fiction, Myth, and Imagination. *American Philosophical Quarterly* 41: 331–337.
- Fine, Kit 1982. The Problem of Non-Existents. I. Internalism., *Topoi* 1: 97–140.
- Kaplan, David 1973. Bob and Carol and Ted and Alice. In K.J.I. Hintikka, J.M.E. Moravcsik, and P. Suppes (eds.) *Approaches to Natural Language*. Dordrecht: Reidel, 490–518.
- Kaplan, David 1989. Afterthoughts. in Joseph Almog *et al.* (eds.) *Themes from Kaplan*. Oxford: Oxford University Press, 565–614.
- Kripke, Saul 1963. Semantical Considerations on Modal Logic. *Acta Philosophica Fennica* 16, 83–94.
- Kripke, Saul 1971. Identity and Necessity. In Milton K. Munitz (ed.) *Identity and Individuation*. New York: New York University Press, 135–164.
- Kripke, Saul 1972. Naming and Necessity. In Donald Davidson and Gilbert Harman (eds.) *Semantics of Natural Language*, 2nd edition. Dordrecht: D. Reidel, 253–355; Reprinted with Preface, *Naming and Necessity*. Cambridge, MA: Harvard, 1980, page references are to the reprint.
- Kripke, Saul 1973. Reference and Existence (The John Locke Lectures), The Saul Kripke Center Archives, The CUNY Graduate Center. (Copies available from The Saul Kripke Center upon request.)
- Lewis, David 1973. *Counterfactuals*. Oxford: Blackwell.
- Lewis, David 1978. Truth in Fiction. *American Philosophical Quarterly* 15: 37–46. Reprinted in David Lewis, *Philosophical Papers* (Volume 1), Oxford: Oxford University Press, 1983, 261–275.
- Lewis, David. 1986. *The Plurality of Worlds*. Oxford: Blackwell.
- Mally, Ernst 1912. *Gegenstandstheoretische Grundlagen der Logik und Logistik*. Leipzig: Barth.
- Meinong, Alexius 1904. Über Gegenstandstheorie. In Alexius Meinong (ed.) *Untersuchungen zur Gegenstandstheorie und Psychologie*, Leipzig: Barth. Translation by I. Levi, D.B. Terrell, and R.M. Chisholm, in Roderick Chisholm (ed.) *Realism and the Background of Phenomenology*. Glencoe: Free Press, 1960, 76–117.
- Parsons, Terence 1980. *Nonexistent Objects*. New Haven: Yale University Press.

<sup>36</sup> For comments on the talk version of this paper, I'd like to thank participants and organizers of the conference *Realism within Phenomenology and within Analytic Philosophy* held at Kaposvár University in January 2012. Special thanks are due to András Simonyi, Zsolt Kapelner and an anonymous referee for many thoughtful and incisive suggestions on an earlier draft of this paper.

- Rosen, Gideon 2012. Abstract Objects. In Edward N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Spring 2012 Edition), URL = <<http://plato.stanford.edu/archives/spr2012/entries/abstract-objects/>>.
- Sainsbury, Mark 2010. *Fiction and Fictionalism*. London: Routledge.
- Salmon, Nathan 1998. Nonexistence. *Noûs* 32, 277–319.
- Searle, John R. 1979. The Logical Status of Fictional Discourse. In P.A. French *et al.* (eds.) *Contemporary Perspectives in the Philosophy of Language*. Minneapolis: University of Minneapolis Press, 233–243.
- Schiffer, Stephen 1996. Language-Created Language-Independent Entities. *Philosophical Topics*, 24, 149–166.
- Thomasson, Amie L. 1999. *Fiction and Metaphysics*. Cambridge: Cambridge University Press.
- Thomasson Amie L. 2009. Fictional Entities. In Jaegwon Kim, Ernest Sosa, Gary Rosenkranz (eds.) *A Companion to Metaphysics*, Second edition. Oxford: Blackwell, 10–18.
- Van Inwagen, Peter 1977. Creatures of Fiction. *American Philosophical Quarterly* 14: 299–308.
- Van Inwagen, Peter 2000. Quantification and Fictional Discourse. In Anthony Everett – Thomas Hofweber (eds.) *Empty Names, Fiction, and the Puzzles of Nonexistence*. Stanford: CSLI Publications, 235–247.
- Walton, Kendall L. 1990. *Mimesis as Make-Believe*. Cambridge, MA: Harvard University Press.
- Yagisawa, Takashi 2001. Against Creationism in Fiction. *Philosophical Perspectives* 15: 153–172.
- Zalta, E.N. (1983). Abstract Objects: An Introduction to Axiomatic Metaphysics, Dordrecht: D. Reidel.
- Zalta, Edward N. 2000. Pretense Theory and Abstract Object Theory. In Anthony Everett – Thomas Hofweber (eds.) *Empty Names, Fiction, and the Puzzles of Nonexistence*. Stanford: CSLI Publications, 117–147.
- Zalta, Edward N. 2006. Deriving and Validating Kripkean Claims Using the Theory of Abstract Objects. *Noûs* 40/4: 591–622.
- Zvolenszky, Zsófia 2007. Naming with Necessity. In *Modality, Names and Descriptions*. Ph.D. thesis, New York University.