

Counting experiments

Jonathan Livengood¹

Published online: 27 November 2017

© Springer Science+Business Media B.V., part of Springer Nature 2017

Abstract In this paper, I show how one might resist two influential arguments for the Likelihood Principle by appealing to the ontological significance of creative intentions. The first argument for the Likelihood Principle that I consider is the argument from intentions. After clarifying the argument, I show how the key premiss in the argument may be resisted by maintaining that creative intentions sometimes independently matter to what experiments exist. The second argument that I consider is Gendenberger’s (Br J Philos Sci 66(3):475–503, 2015) rehabilitation of Birnbaum’s (J Am Stat Assoc 57(298):269–306, 1962) proof of the Likelihood Principle from the (supposedly) more intuitively obvious principles of conditionality and sufficiency. As with the argument from intentions, I show how Gendenberger’s argument for his Experimental Conditionality Principle may be resisted by maintaining that creative intentions sometimes independently matter to what experiments exist.

Keywords Likelihood Principle · Analytic metaphysics · Experiments · Evidential value · Creative intentions · Artifacts · Abstract creationism · Conditionality principle · Arbitrariness

Suppose I want to know whether the proportion of likely voters who support the liberal candidate in an upcoming local election is statistically different from fifty percent. I go out to poll likely voters, canvassing a few randomly chosen neighborhoods in order to get a representative sample. Here are two ways I might proceed: I could poll 50 likely voters and then stop and record the number who

✉ Jonathan Livengood
jlive@illinois.edu

¹ University of Illinois at Urbana-Champaign, Urbana, IL, USA

support the liberal candidate; or I could poll likely voters until I find 19 who support the liberal candidate and then stop and record the number polled. It is well-known among philosophers of statistics that such *stopping rules* sometimes matter for the conclusions that a frequentist should draw from an experiment.¹ And it is well known that stopping rules in such cases should *not* matter to anyone endorsing the *Likelihood Principle*.²

Perhaps stopping rules should not matter. But if so, it is non-obvious. As Savage remarks (1962, 18), the evidential irrelevance of stopping rules flies in the face of long statistical tradition. So, why endorse the Likelihood Principle or its implications for stopping rules? In this paper, I consider two arguments offered by philosophers and statisticians in support of the Likelihood Principle, and I show how both arguments may be resisted by maintaining that creative intentions sometimes independently matter to what experiments exist.

1 Creative intentions and the ontology of experiments

The first argument that I will consider for the Likelihood Principle has come to be called *the argument from intentions*. In discussing the argument from intentions as he heard it from Barnhard, Savage (1962, 76) puts it as follows: “The design of a sequential experiment is, in the last analysis, what the experimenter actually intended to do. His intention is locked up inside his head and cannot be known to those who have to judge the experiment.” From the fact (if it is a fact) that the experimenter’s intention is locked up inside his head, it is supposed to follow that

¹ A frequentist who uses a binomial model for the sampling problem here and who assumes the standard 0.05 significance level for a test of the hypothesis that the true proportion is fifty percent will reject the hypothesis when using the rule “stop at 19 successes” ($p = 0.0427$). But such a frequentist will fail to reject the hypothesis when using the rule “stop at 50 total observations” ($p = 0.0595$). The canonical example is due to Savage (1962, 18).

² The Likelihood Principle may be formulated in several different ways. Berger and Wolpert (1988, 1) begin their monograph on the Likelihood Principle as follows: “Among all prescriptions for statistical behavior, the Likelihood Principle (LP) stands out as the simplest and yet most far-reaching. It essentially states that all evidence, which is obtained from an experiment, about an unknown quantity θ , is contained in the likelihood function of θ for the given data. The implications of this are profound, since most non-Bayesian approaches to statistics and indeed most standard statistical measures of evidence (such as coverage probability, error probabilities, significance level, frequentist risk, etc.) are then contraindicated.” Edwards (1972, 30) states the Likelihood Principle this way: “Within the framework of a statistical model, *all* the information which the data provide concerning the relative merits of two hypotheses is contained in the likelihood ratio of those hypotheses on the data.” Gandenberger (2015, 476–477) says that according to the Likelihood Principle, “two experimental outcomes are evidentially equivalent if they have proportional likelihood functions—that is, if the probabilities that the set of hypotheses under consideration assign to those outcomes are proportional as functions of those hypotheses.”

The term “experiment” as I use it in this paper and as it is widely used in probability theory, such as in the Berger and Wolpert quotation in this footnote, does not implicate the presence of experimental control but means something closer to “empirical observation” or perhaps “empirical test of a conjecture.” As Gandenberger (2015, 478) remarks, “This broad use of the term ‘experiment’ is not ideal, but there is no alternative that is obviously better.”

stopping rules do not have any evidential import and that the Likelihood Principle is correct.

Mayo (1996, 346–350) criticizes the argument on the grounds that intentions matter for all (or nearly all) of the properties an experiment has. She writes (347): “Any and all aspects of what goes into specifying an experiment could be said to reflect intentions—sample size, space of hypotheses, prediction to test, and so on—but it does not mean that paying attention to those specifications is tantamount to paying attention to the experimenter’s intentions.” However, Mayo’s criticism seems to miss an important difference between stopping rules and sample sizes (and the like). A stopping rule *just is* an experimenter’s intention, but a sample size is a property that an experiment has *in virtue of* an experimenter’s intention. Writing it out explicitly, I take the argument from intentions to run something like this:

- [A1] If two experiments differ only with respect to their stopping rule, then they differ only with respect to what some experimenters intended to do.³
- [A2] If two experiments differ only with respect to what some experimenters intended to do, then they do not differ in evidential value.⁴

- [A3] If two experiments differ only with respect to their stopping rule, then they do not differ in evidential value.

I take Mayo to be criticizing [A2]. But if so, the criticism misfires, since two experiments that differ *only* with respect to, say, an experimenter’s intention to choose a sample of a given size do not differ with respect to the size of the sample actually drawn. The sample size depends on some intention—or at least, it typically does. But the property that matters evidentially is the sample size, not the intention to draw a sample of that size. Hence, there is a relevant difference between properties such as an experiment’s sample size on the one hand and an experiment’s stopping rule on the other. The sample size, but not the stopping rule, is a property

³ One might worry that this premiss assumes too much and is in fact inconsistent with what proponents of the Likelihood Principle accept. After all, Howson and Urbach (1996, 214) say that stopping rules *do* matter in some cases. Later in their book, they give an example of what they take to be an informative stopping rule and then remark (366): “Hence, the stopping rule is not necessarily uninformative, but as this example suggests, normally it would be. This concession should not be misunderstood. It does not mean that the scientist’s *intention* to stop the trial at a particular point is of any inductive significance.” With respect to my reconstruction of the argument from intentions, Howson and Urbach are being too charitable. A pair of experiments that differ with respect to their stopping rules in the manner suggested by Howson and Urbach’s example also differ with respect to the distribution of individuals in a population being sampled. And it is the latter difference that matters for those who endorse the Likelihood Principle. Hence, Howson and Urbach’s remarks are no threat to premiss [A1] in my reconstruction.

⁴ The notion of *evidential value* needs to be treated carefully in order to allow for the possibility that different epistemic agents might draw different conclusions from experiments with identical evidential value. For example, the agents might have different initial credences or they might deploy different rules of inference. But for present purposes, common sense should suffice.

with a life of its own. Or at least, a proponent of the argument from intentions might maintain as much.

So, [A2] seems to survive Mayo's criticism. Even so, one might feel the need for some further argument in favor of [A2]. Other authors have used decision theoretic, statistical, and simulational tools in trying to resolve the debate about stopping rules.⁵ In what follows, I consider an alternative metaphysical approach inspired by puzzles about co-located objects. I first lay out an argument for [A2]. I then show how one might resist the key premiss in the argument for [A2] if one accepts that creative intentions sometimes independently matter to what experiments exist. Tracking the argument back leads into some discussion of the ontology of experiments.

Consider the following variation on the story from the beginning of this paper.⁶ For later reference, call the first story S_1 and the new variant S_2 . Suppose that my long-time collaborator Sally and I want to know whether the proportion of likely voters who support the liberal candidate in an upcoming local election is statistically different from fifty percent. Further suppose that Sally and I share all of our initial opinions.⁷ We go out together to poll likely voters, canvassing a few randomly chosen neighborhoods in order to get a representative sample. I carry a clipboard, Sally asks the questions, and I write down the responses. But suppose that Sally and I did not talk about our sampling plan in advance and had different procedures in mind. Sally intended to poll 50 likely voters and then record the number who support the liberal candidate, while I intended to poll as many likely voters as needed until finding 19 who support the liberal candidate and then record the total number of likely voters polled. And now suppose that by coincidence, Sally and I are jointly satisfied with our work and stop at the same time: the 50th person interviewed just happened to be the 19th to say that she supports the liberal candidate.

How many experiments did Sally and I perform in S_2 ? An obvious and natural answer is that we conducted exactly one experiment. By construction, we asked the same questions to the same people on the same day. We kept a single record of the responses. And so on. It is true that Sally and I had different intentions when we set out, but the difference in our intentions didn't matter in the actual case, since the sample we actually drew satisfied both Sally's intentions and mine. Intentions on their own are not ontologically significant. These considerations suggest the following argument for premiss [A2] in the argument from intentions.

⁵ For a few examples, see Mayo and Kruse (2001), Sprenger (2009), Steele (2013), Yu et al. (2014), and Rouder (2014).

⁶ Howson and Urbach (1996, 212) tell a similar story. However, they do not draw attention to the issues that I care about in this paper.

⁷ By saying that Sally and I have all the same initial opinions, I mean to imply at least this much: that we have exactly the same credences and exactly the same credence update rules. That Sally and I share all of our initial opinions will be an important stipulation in the arguments considered later on.

- [B1] Intentions cannot matter to what experiments exist.
 [B2] If intentions cannot matter to what experiments exist and two experiments differ only with respect to what some experimenters intended to do, then those experiments are identical.
-

- [B3] If two experiments differ only with respect to what some experimenters intended to do, then those experiments are identical.
 [B4] If two experiments are identical, then they do not differ in evidential value.
-

- [A2] If two experiments differ only with respect to what some experimenters intended to do, then they do not differ in evidential value.

The argument is valid, so if one wants to reject the conclusion, one should reject at least one of the premisses. Suitably regimented, premiss [B2] is a logical truth. So, it should be uncontroversial. Similarly, premiss [B4] appears unassailable: properly understood, it follows from the indiscernibility of identicals. What about premiss [B1]?

One might initially think that intentions can *clearly* matter to what experiments exist. After all, Galileo's experiments with inclined planes, Newton's experiments with prisms, and *thousands* of other experiments would never have been conducted at all without an experimenter intending to investigate a specific question, intending to set up equipment in a specific way, and so on. The objection here may be deflected by slightly amending premiss [B1] to say that intentions cannot *independently* matter to what experiments exist and then adjusting the rest of the argument accordingly. The thought here is that intentions have no ontological significance above and beyond contributing in whatever way they contribute to the actions by which one arranges the basic furniture of the world. If being in some specific arrangement is not sufficient for some *xs* to compose an experiment, then the *xs* will not compose an experiment even if an agent intends to bring an experiment into existence by arranging the *xs* in that specific way and succeeds in arranging them in just the way she intended. And if being in some specific arrangement *is* sufficient for some *xs* to compose an experiment, then what the experimenter intends to do makes no difference to what there is. In other words, the ontological question is completely settled by how things are arranged: intentions contribute nothing further.

According to the slightly amended version of premiss [B1], which I will call [B1*], intentions cannot *independently* matter to what experiments exist. But one might resist [B1*] by maintaining both that intentions can independently matter to what artifacts exist and also that experiments are artifacts:

- [D1] If experiments are artifacts, then intentions can independently matter to what experiments exist.
 [D2] Experiments are artifacts.
-

[D3] Intentions can independently matter to what experiments exist.

How good are the premisses? Experiments have many hallmarks of artifacts, which one may confirm by considering some prominent theories of what makes something an artifact.⁸ For example, according to Hilpinen (1993, 156), something is an artifact if and only if it has an author. Experiments have authors (in Hilpinen's broad sense of "author"). There are Boyle's experiments with the air pump, Lavoisier's experiments on combustion, Curie's experiments with uranium minerals, Milgram's experiments on obedience to authority, and so on.

Baker (2007, 49) tells us that artifacts include "everything that is produced intentionally—paintings and sculptures as well as scissors and microscopes." Thomasson (2003, 2007) defends the view that an artifact must be the product of someone's intention to produce something of the kind to which it belongs. Irmak (2013) explicitly defines artifacts to be intentional products of human activity, and he remarks (in Footnote 2) that his definition is "widely accepted." Experiments are intentional products of human activity that are intended to be in the kind "experiment" and are often intended to be in a narrower kind, such as "experiment to test hypothesis *H*." Experiments are designed, they (typically) need to be conducted carefully in order to have interesting results, and they are often modified in order to investigate new questions or to control for previously unforeseen confounds, all of which are marks of intentionality, productivity, or both.

Houkes and Vermaas (2009) offer a more complicated account, according to which an artifact is an item that is "created by a successful execution of a make plan" (414), where a make plan is a list of actions—such as welding, cutting, bending, and so on—involved in producing an item. They emphasize the distributed nature of contemporary production of artifacts, where separate teams of individuals may be responsible for what happens at various stages from design to manufacture. But the features of contemporary large-scale production of artifacts that Houkes and Vermaas use to undermine what they call the *artisan model* are also present in contemporary large-scale scientific experiments. On the extreme end, there are experiments involving the collision of high-energy particles at the Large Hadron Collider, such as the TOTEM experiment (Anelli et al. 2008) and the ATLAS experiment (Aad et al. 2008), which require the collaboration of dozens or even hundreds of researchers. Even simpler experiments today (in disciplines ranging from materials science to psychology) often have dedicated statisticians proposing an experimental design, area experts developing instruments for the specific setting of interest, and then laboratory researchers bringing the experiments to completion. Hence, experiments are artifacts according to the account of Houkes and Vermaas as well.

Hence, the accounts of artifacts endorsed by Houkes and Vermaas, Baker, Thomasson, and Irmak all lend support to premiss [D2] in the argument, and no

⁸ One objection that may occur to the reader is the possibility of so-called *natural experiments*. Plausibly, one may respond to the objection by treating a natural experiment as a found object that is repurposed in some way, like a piece of driftwood that is used, without modification, as a wine rack. See Korman (2015, 155–156) for discussion.

account of artifacts maintains that experiments are *not* artifacts.⁹ Moreover, every account of artifacts is at least consistent with the view that if an experiment is an artifact, then creative intentions can independently matter to what experiments exist. But consistency is a low bar, and one might reasonably want some positive reason for thinking that [D1] is true. One (admittedly weak) reason to think that [D1] is true is that the naïve view appears to be that artifacts are what they are in part as a result of creative intentions.¹⁰ For example, ordinary artifact-categorization judgments are independently influenced by being told that a designer intended the artifact to have a specific function. Another reason to think that [D1] is true is suggested by Baker's claim (2007, 211) that being a statue is more than having this or that arrangement of parts. As she writes, "Atoms arranged statuesquely are not (and do not constitute) a statue in a world lacking artists and the conventions of art. A meteor that looks like a statue is not a statue." Reflecting on the case of a meteor that looks like a statue, Korman explains the difference by appeal to creative intentions, writing (2015, 153):

Creative intentions are indeed relevant to which kinds of things there are. Suppose that a meteoroid, as a result of random collisions with space junk, temporarily comes to be a qualitative duplicate of some actual statue. Intuitively, nothing new comes into existence which, unlike the meteoroid, cannot survive further collisions that deprive the meteoroid of its statuesque form. Likewise, unintentionally and momentarily kneading some clay into the shape of a gollyswoggle does not suffice for the creation of something that has that shape essentially.... The fact that many have set out to make statues, while no one has ever set out to make a gollyswoggle, is an ontologically significant difference between statues and gollyswoggles.

To the extent that one agrees with Baker and Korman about the meteor case, there is some (defeasible) reason to think that creative intentions have independent ontological significance for artifacts. But these points are not peculiar to meteors and statues, they apply to artifacts in general. Hence, in the absence of some reason to think that experiments form an unusual kind of artifact, anyone who accepts that intentions can independently matter to what *artifacts* exist ought to accept that intentions can independently matter to what *experiments* exist.

⁹ Philosophers have also debated what it is that makes something a member of a specific artifactual kind Hilpinen (2011). In addition to already-cited papers by Thomasson, see Schwartz (1978) and Kornblith (1980). Taking a page from the metaphysics of personal identity and personhood (especially Thomson 2008), one might hope that the features—whatever they are—that serve to distinguish different kinds of artifact will have a close connection to what makes something an artifact in the first place. For example, Kornblith (1980, 112) writes, "At least for the most part, it seems that what makes two artifacts members of the same kind is that they perform the same function." If so, one might hope that *having a function* is part of the correct account of what it is to be an artifact. But again, experiments have functions, so accounts that treat function as central to being an artifact should count experiments as artifacts.

¹⁰ The claim that creative intentions matter according to the naïve view is supported by psychological work on mereological composition (for which, see Rose and Schaffer 2015, and Korman and Carmichael 2017) and on artifact categorization judgments (for which, see Bloom 1996; Malt and Sloman 2007; Barrett et al. 2008; Chaigneau et al. 2008; Chaigneau et al. 2016; and Volume 4, Issue 3 of the *Review of Philosophy and Psychology*).

Return now to the question: How many experiments did Sally and I perform in S_2 ? The answer “one” suggested an argument for [A2], which is itself a premiss in an argument against the evidential relevance of stopping rules. By extension, the argument from intentions is supposed to give us reason to endorse the Likelihood Principle and to reject standard frequentist statistical practices. But one might reject [B1] in the argument for [A2] on the grounds that experiments are artifacts: exactly the sorts of things for which creative intentions have independent ontological significance. Reflecting on the puzzle of the statue and the clay, a philosopher might say that Sally and I conducted two different experiments in S_2 , and those experiments just happen to be co-located. After all, Sally’s experiment and my experiment have different modal profiles. Sally’s experiment could not have included more than 50 likely voters in total, but it could have included more than 19 likely voters who support the liberal candidate. By contrast, my experiment could not have included more than 19 likely voters who support the liberal candidate, but it could have included more than 50 likely voters in total. Hence (by the indiscernibility of identicals), Sally’s experiment is not identical to mine. Rejecting [B1] then does work by underwriting an explanation of the differences between Sally’s experiment and mine in terms of the different intentions of the experimenters.

A proponent of the argument from intentions against the evidential relevance of stopping rules might try to deny [D2] by claiming that experiments are abstract objects. The underlying thought here, which I will label [NAA] for future reference, is that no abstract object is an artifact. Here are three relatively quick arguments one might give (and that some philosophers have given) in defense of [NAA]. First, one might argue that [E1] abstract objects are eternal—neither coming into existence nor going out of existence—but that [E2] artifacts come into existence when they are created. So, no abstract object is an artifact. Second, one might argue that [F1] artifacts are sensitive to causal influences, but [F2] abstract objects are not. So again, no abstract object is an artifact. Third, one might argue that [G1] abstract objects are not located anywhere in space but that [G2] artifacts are located in space. Hence, no abstract object is an artifact.

Now, if no abstract object is an artifact, then in order to deny that experiments are artifacts, one only needs to show that experiments are abstract objects. One reason to think that experiments are abstract objects is that experiments are repeatable. Just as one may play Beethoven’s *Für Elise* if one has a piano, one may conduct Mendel’s genetics experiments if one has a garden. And just as *Für Elise* is not anything concrete, such as Beethoven’s original score or the sounds produced on any particular occasion when a pianist plays through the score, Mendel’s genetics experiments are not concrete either. One might go further and say that it is an essential feature of experiments that they be repeatable. Many people take science to trade fundamentally in what may be tested and observed over and over again. If experiments were not repeatable, it would seriously undermine the common view that science proceeds on independently verifiable experimental evidence. Moreover, we often become sure of an experimental result only after it has been replicated. Hence, one might argue that experiments are repeatable and that if they are repeatable, they are abstract objects. If experiments are abstract objects and if no

abstract object is an artifact, then [D2] is false: experiments are not artifacts. Hence, a proponent of the argument from intentions might try to defend [B1*] by maintaining that experiments are abstract objects.¹¹

Such a view of experiments is akin to the Platonist theory of musical works defended by Dodd (2000, 2002, 2004, 2007). According to Dodd (2004, 342), a musical work is a performance type: “an abstract entity whose identity is determined by the condition a particular must meet, if it is to count as one of its tokens.”¹² Hence, for Dodd, a musical work is the sort of thing that is discovered, not created. However, many philosophers who have considered the matter reject the Platonist approach to music in part because they find it deeply counterintuitive to say that musical works are not created. As Levinson (1980, 8) puts it in rejecting the view that musical works are pure sound structures:

There is probably no idea more central to thought about art than that it is an activity in which participants create things—these things being artworks. The whole tradition of art assumes art is creative in the strict sense, that it is a godlike activity in which the artist brings into being what did not exist beforehand—much as a demiurge forms a world out of inchoate matter. ... The suggestion that some artists, composers in particular, instead merely *discover* or *select* for attention entities they have no hand in creating is so

¹¹ At this point, one might wonder whether Platonism with respect to experiments offers any *positive* reason for thinking that [B1*] is true or whether it merely cuts off an argument *against* [B1*]. Plausibly, an experimentalist’s intentions matter with respect to which experiment she selects to perform. But that isn’t enough to secure the claim that intentions have independent ontological significance. There are two separable questions here. First, one might wonder whether the intentions that matter with respect to which experiment one selects to perform ever matter independently with respect to which experiment one *actually performs*. If the identity of the experiment one performs supervenes in the right way on what one actually does, then the intention to select one experiment rather than another will not matter independently. One plausible consequence of such a view is that an experimentalist might perform an experiment different from the one she selects to perform. For example, a male researcher might intend to replicate an experiment on some laboratory rats first conducted by a female colleague but fail to actually replicate the experiment because he is unaware of the fact that male experimenters cause rodents to experience higher stress levels (for which, see Sorge et al. 2014). He selects the same experiment to perform, perhaps borrowing “aboutness” in the way Pollard (2007) suggests we borrow reference in mathematical discourse. But he does not *perform* the same experiment. (An interesting question here is how to distinguish between performing a different experiment correctly and performing the same experiment incorrectly.) Second, one might wonder whether there is any experiment that simply could not be performed—even accidentally—without having some specific intentions regarding *how* to perform it. A stopping rule is an intention regarding how to conduct an experiment, not an intention about what experiment to conduct. But it is hard to see how such an intention could be *necessary*. For any case in which an experimentalist intends to sample in a given way, it seems that we can imagine an experimentalist who samples in the same way without reflecting on her experimental design. For example, we might imagine an experimentalist who samples until she gets bored.

Returning to the referee’s question: I am not sure whether Platonism about experiments provides any positive reason for thinking that [B1*] is true. I am, at a gut level, inclined to think that Platonism about experiments favors proponents of the Likelihood Principle. But I have not been able to produce any arguments for the claim that Platonism about experiments entails [B1*] that I find satisfying.

¹² In earlier writings, Dodd says that a musical work is a type of sound-sequence-occurrence, but his later writing implicates (at least) that he has not changed his view: sound-sequence-occurrences are performances.

contrary to this basic intuition regarding artists and their works that we have a strong *prima facie* reason to reject it if we can.

To set his target clearly, Dodd (2000, 426) lays out the following *argument from creatability* against the “most natural and common proposal” regarding the ontology of musical works: namely that musical works are sound structures¹³:

- [H1] Sound structures exist at all times.
- [H2] If musical works were sound structures, they could not be created (that is, brought into being) by their composers.
- [H3] But musical works *are* created by their composers.
- [H4] Musical works are not sound structures.

With a few minor modifications, we may transform the argument from creatability into an argument against the claim that experiments are abstract objects. So, it may be useful to see how Dodd resists the argument. Dodd denies [H3], but he does not agree that doing so is counterintuitive. Of course, he agrees that it would be counterintuitive to deny that composers are *creative*, but he thinks it is possible to be creative without having any god-like power of creation. He writes (2000, 428):

A composer is creative, not through bringing works into existence, but by having to exercise imagination in composing the works she does.... A creative composer is someone who has the imagination to compose works of music that others do not have the capacity to compose. Composition is, indeed, a form of discovery; but discoveries can be creative.

According to Dodd, proponents of the argument from creatability have found [H3] plausible because they have confused *being creative* with *having a god-like power of creation*.¹⁴

At this point, we seem to face a difficult choice. On one hand, we seem to have good reason to think that experiments are artifacts. Assuming the principle [NAA] that no abstract object is an artifact, we may then infer that experiments are not abstract objects. On the other hand, we seem to have good reason to think that experiments are abstract objects. Again, assuming the principle [NAA] that no abstract object is an artifact, we may infer that experiments are not artifacts. Clearly, one may either reject the claim that experiments are artifacts or reject the claim that experiments are abstract objects. But there is also a third option one might try: reject the principle [NAA] that no abstract object is an artifact. Rejecting [NAA] strikes me as the correct move here. The arguments that I rehearsed for [NAA] seem both

¹³ I am quoting Dodd here except for relabeling the propositions. In doing so, I am adopting his use of “so” instead of using a line representing “therefore” as in previous arguments.

¹⁴ Deutsch (1991) provides an alternative approach to resisting the argument from creatability: reject the inference from [H1] to [H2]. Deutsch argues that the creation of a melody (a musical work) or a story (a work of fiction) does not require bringing anything into existence. He writes (221): “To be in a position to (literally) create a thing is to be in a position to stipulate rather than merely describe what the thing is like.” Deutsch’s paper is very interesting, but at the end of the day, I agree with Brock (2010, 343) that “*Deutschian* creation isn’t a kind of creation at all; causal creation is the only variety of genuine creation.” And hence, I think that Deutsch’s strategy is not a good one.

weaker than the reasons we have for thinking that experiments are abstract objects and also weaker than the reasons we have for thinking that experiments are artifacts.¹⁵

Thus, opponents of the argument from intentions may reasonably reply to the claim that experiments are abstract objects by agreeing: experiments *are* abstract objects; they are *also* artifacts. Rejecting [NAA], the argument that intentions can independently matter to what experiments exist may be slightly amended as follows:

[D1*] If experiments are abstract artifacts, then intentions can independently matter to what experiments exist.

[D2*] Experiments are abstract artifacts.

[D3] Intentions can independently matter to what experiments exist.

Rejecting [NAA] is neither unmotivated nor ad hoc. Rather, denying [NAA] is a standard move in the literature on fictional characters, and abstract creationism has been defended, elaborated, or suggested as an obvious example with respect to musical works, software, laws, words, games, recipes, traditions, and brand-names.^{16,17} Since abstract artifacts are the sorts of things that depend on creative intentions for their very existence, premiss [D1*] is, if anything, even more plausible than premiss [D1]. Hence, endorsing abstract creationism with respect to experiments appears to be a promising way to resist the argument from intentions for the Likelihood Principle.¹⁸

At this point, I want to step back, take stock, and answer an objection from an anonymous referee who points out that one might have a good argument to the effect that two experiments are metaphysically distinct and yet feel no pressure to

¹⁵ Since I think that the best option at this point in the dialectic is to reject [NAA], I have to dispute at least one premiss in each of the three arguments I suggested earlier in support of [NAA]. In the first argument, I deny the premiss [E1] that abstract objects are eternal. In the second argument, I deny the premiss [F2] that abstract objects are not sensitive to causal influences. In the third argument, I lean toward rejecting [G2] and saying that while experiments and other abstract artifacts are not located in space, they are still artifacts. However, I sometimes think that there are abstract objects that do have spatial locations. For example, one might think that a set of concrete objects is itself an abstract object that is located in the fusion of the spaces occupied by its members. If so, then I could retain [G2] and understand experiments to be spatially-located abstract artifacts. But along this route, I worry that the term “abstract object” becomes too mysterious—too unmoored from its history to do any work.

¹⁶ Brock (2010, 339) writes, “Creationism about fictional characters seems to be the orthodoxy in philosophical circles today.” He then goes on to quote extensively from contemporary philosophers endorsing the orthodox view. These include van Inwagen (1977), Salmon (1998), Thomasson (1999), and Soames (2002).

¹⁷ For musical works, see Levinson (1980) and discussion in Caplan and Matheson (2004); for software, see Irmak (2013); for laws, see Burazin (2016); for words, see Sainsbury and Tye (2012); for games, recipes, and the rest, see lists in Korman (2014) and in Zvolenszky (2012).

¹⁸ Of course, criticisms have been raised against abstract creationism. However, criticism of abstract creationism has (as far as I know) focused mainly on its application to fictional characters, and fictional characters may very well behave differently from experiments and other alleged examples of abstract artifacts. For two very different and interesting criticisms of abstract creationism with respect to fictional characters, see Brock (2010) and Vecsey (2014). See Friedell (2016) for a response to Brock.

say that they have different evidential values. The referee writes, “Even if there’s a good metaphysical argument that an experiment performed while wearing a red hat is metaphysically distinct from an otherwise identical experiment performed while wearing a blue hat, there’s nothing at all unnatural about claiming that that metaphysical difference is epistemologically irrelevant.”¹⁹ Indeed, nothing I have said goes to show that two different experiments always have different evidential value or that two experiments that differ only in the intentions of some experimenters will (or must) have different evidential values. And that is a very good thing, since it would be *false* to say that different experiments always have different evidential value. Different experiments *sometimes* have different evidential values, but they do not *always* have different evidential values.

What physical or metaphysical differences between experiments make an epistemological difference? According to the Likelihood Principle, a difference in stopping rules is not enough to make an epistemological difference. But why this should be the case is not obvious. Rather, it seems obvious (at least to me) that stopping rules *should* matter evidentially. Of course, even if it seems obvious to you that stopping rules should matter, you might be convinced by argument to set aside your intuitions and endorse the Likelihood Principle. One possible argument in support of the Likelihood Principle rests on the plausible-looking but undefended assumption [A2] that if two experiments differ only with respect to what some experimenters intended to do, then those experiments do not differ in evidential value. At this point, I suggested an argument for [A2] that makes use of the assumption [B1*] that intentions cannot independently matter to what experiments exist. The argument from [B1*] to [A2] fails if creative intentions sometimes independently matter to what experiments exist.

Perhaps any differences in the creative intentions that two experimenters have are as epistemologically uninteresting as differences in the color of their hats. However, we have so far seen only one *argument* for that claim—one that rests on the premiss that a mere difference in the creative intentions of some experimenters is not enough to make the experiments they perform distinct. That argument may be resisted by showing that a difference in the intentions of two experimenters sometimes makes a difference to the identity of the experiments they perform and thus denying its crucial premiss. Consider an analogous case in the referee’s colored hats scenario. Suppose that one were arguing for the claim that the hat color of an experimenter makes no epistemological difference and that one appealed to the claim that the hat color of an experimentalist makes no difference to the identity of the experiment being performed. Then one could resist the conclusion of the argument by showing that some differences in the hat color of an experimentalist *do* make a difference to the identity of the experiment being performed.

¹⁹ I do not want to dwell on them, but there are some disanalogies between the referee’s hat case and cases involving experimenters with different intentions. First, no one initially thinks that hat color is epistemically relevant. But many people have thought and still think that stopping rules *are* epistemically relevant. Second, it seems initially plausible that hat color is *metaphysically* relevant to what experiment one performs. But the opposite is initially plausible with respect to creative intentions.

Now, perhaps the argument from intentions works in some other way. Maybe it depends on some plausible principle(s) that I have overlooked. But if so, proponents of the Likelihood Principle will need to articulate the argument and defend its premisses. Alternatively one might give up on the argument from intentions and argue directly for the Likelihood Principle in a different way. In the next section, I consider such an approach for defending the Likelihood Principle: derive it from (apparently more secure) axioms. I show how the assumption that creative intentions are ontologically significant, which I showed could be used to resist the argument from intentions, may be called on to resist the axiomatic argument as well.

2 Creative intentions and conditionality principles

Some statisticians and philosophers have sought to defend the Likelihood Principle by deriving it from seemingly more obvious starting points. In this section, I consider Gandenberger's (2015) proof of the Likelihood Principle, and I show how asserting the ontological significance of creative intentions provides a principled way to reject the Experimental Conditionality Principle (ECP), which is one of two assumptions Gandenberger calls on in his proof of the Likelihood Principle. I begin by discussing the thought experiment that Gandenberger calls on to support the ECP. I suggest a way of using the thought experiment to support one (crucial) premiss in an arbitrariness argument for the (comparatively modest) claim that the ECP holds in cases like S_1 . I then recommend a strategy for resisting the argument, which makes use of the now-familiar claim that creative intentions are ontologically significant. My thinking here is indebted to Korman's (2015) discussion of the arbitrariness argument against Conservatism with respect to ordinary objects and especially by his responses to various arbitrariness arguments.²⁰ Having addressed the arbitrariness argument, I claim that philosophers who endorse the claim that creative intentions are ontologically significant ought to say that Gandenberger has misidentified the intuition being pumped by his thought experiment. Such philosophers should say that the salient intuition is a metaphysical one having to do with what experiment was actually performed, rather than an epistemological one having to do with the evidential value of a peculiar experimental design.

In order to state Gandenberger's Experimental Conditionality Principle (ECP), we need the idea of a mixture experiment. A *mixture experiment* is an experiment in which a random process is used to select—from a collection of *component* experiments—one experiment to actually perform. For example, an experiment in which one rolls a die in order to determine which of six different questions to ask in conducting a survey is a mixture experiment. With the idea of a mixture experiment in hand, the ECP may be stated informally as follows: The outcome of a mixture experiment is evidentially equivalent to the corresponding outcome of the component experiment actually performed. Gandenberger motivates the ECP by

²⁰ Korman discusses arbitrariness arguments in Chapter 8 of his book. He considers arguments having to do with artifacts in Chapter 8, Section 4.

way of a thought experiment. Since the example is crucially important to his case, I quote at length, here:

Suppose you work in a laboratory that contains three thermometers, T_1 , T_2 , and T_3 . All three thermometers produce measurements that are normally distributed about the true temperature being measured. The variance of T_1 s measurements is equal to that of T_2 s but much smaller than that of T_3 s. T_1 belongs to your colleague John, so he always gets to use it. T_2 and T_3 are common lab property, so there are frequent disputes over the use of T_2 . One day, you and another colleague both want to use T_2 , so you toss a fair coin to decide who gets it. You win the toss and take T_2 . That day, you and John happen to be performing identical experiments that involve testing whether the temperature of your respective indistinguishable samples of some substance is greater than 0 °C. John uses T_1 to measure his sample and finds that his result is just statistically significantly different from 0°. John celebrates and begins making plans to publish his result. You use T_2 to measure your sample and happen to measure exactly the same value as John. You celebrate as well and begin to think about how you can beat John to publication. ‘Not so fast,’ John says. ‘Your experiment was different from mine. I was bound to use T_1 all along, whereas you had only a fifty percent chance of using T_2 . You need to include that fact in your calculations. When you do, you’ll find that your result is no longer significant.’ (480–481)

For future reference, call the above example *the thermometer case*. Gendenberger describes the experiment that “you” conduct in the thermometer case as a mixture experiment having two simple experiments as components. One simple experiment that is a component of the mixture experiment is a measurement using thermometer T_2 . The other simple experiment that is a component of the mixture experiment is a measurement using thermometer T_3 . As a result of the coin flip, the first simple experiment is actually conducted.

Gendenberger thinks—and I agree—that in the thermometer case, you and John have the same evidence: the coin flip is evidentially irrelevant. But what, if any, further lesson we should draw from the thermometer case is controversial. For example, Wasserman (2012) rejects inferences from examples like the thermometer case to conditionality principles sufficient to justify the Likelihood Principle. Discussing a generic conditionality principle (CP) from which the Likelihood Principle (LP) follows, Wasserman writes (2012, 2): “The main point is that CP (and hence LP) is bogus. Just because it seems compelling that we should condition on the coin flip in the simple mixture example above, it does not follow that conditioning is always good. Making a leap from a simple, toy example, to a general principle of inference is not justified.” Gendenberger replies by claiming that the ECP is not a generalization from the thermometer case. Rather, the ECP is supported by an intuition that we have about the thermometer case. Gendenberger writes (482): “The purpose of the example is merely to make vivid the intuition that features of experiments that could have been but were not performed are irrelevant to the evidential meaning of the outcome of the experiment that actually was

performed. The intuition the example evokes, rather than the example itself, justifies the principle.”

I find Wasserman’s objection unsatisfying. He doesn’t point to any features of realistic cases—such as cases of survey sampling or medical testing—that would distinguish them from toy examples like the thermometer case. But in the absence of distinguishing features, we have no principled reasons for rejecting a generalization. At the same time, I am not convinced that Gandenberger draws the right lesson from the thermometer case. In order to draw out these points and bring us back to creative intentions, I now want to develop an arbitrariness argument in support of the claim that the ECP holds in cases like S_1 . The argument to be developed is analogous to an argument against Conservatism in ordinary object metaphysics.²¹ I am here borrowing both the numbering and the formulation of the premisses from Korman (2015, 153):

[AR33] There is no ontologically significant difference between statues and gollyswoggles.²²

[AR34] If so, then: if there are statues then there are gollyswoggles.

[AR35] There are no gollyswoggles.

[AR36] There are no statues.

The argument as given supports Eliminativism about ordinary objects insofar as it purports to show that at least some ordinary objects—in this case, statues—do not actually exist. The argument could be slightly modified to support Permissivism about ordinary objects by replacing [AR35] with the seemingly innocent claim that there are statues. The conclusion would then be that there are gollyswoggles as well—a result that is also unfavorable to Conservatism. Korman denies [AR33] in

²¹ Conservatism is the thesis that there are ordinary objects—such as tables, chairs, trees, and dogs—but no extraordinary objects—such as trogs, incars, and snowdiscalls. Hence, Conservatism is a middle position between various forms of Eliminativism—which reject some ordinary objects—and various forms of Permissivism—which accept some extraordinary objects. See Korman (2015, 23–25) for further discussion and references.

²² Korman appeals to creative intentions in order to resist an argument from arbitrariness that he attributes to van Inwagen (1990). Van Inwagen writes (126): “Pick up a lump of clay and knead it into some complicated and arbitrary shape. Call anything essentially of that shape a gollyswoggle. Did you bring a gollyswoggle into existence? I should think that if our sculptor brought a statue into existence, then you brought a gollyswoggle into existence. ‘Statue-shaped’ is a less definite shape predicate than ‘gollyswoggle-shaped’, and one we have a use for, and our sculptor intended to produce something statue-shaped while you, presumably, did not intend to produce anything gollyswoggle-shaped. But these facts would seem to be irrelevant to any questions about the existence of the thing produced; if you can make a statue on purpose by kneading clay, then you can make a gollyswoggle by accident by kneading clay. But if you can make a gollyswoggle by accident by kneading clay, then you must, as you idly work the clay in your fingers, be causing the generation and corruption of the members of a compact series of objects of infinitesimal duration. That is what seems to me to be incredible.”

In reply Korman writes (153): “The fact that many have set out to make statues, while no one has ever set out to make a gollyswoggle, is an ontologically significant difference between statues and gollyswoggles, and thus the differential treatment is not arbitrary.” However, following Saenz (2015), Korman does not appeal to creative intentions to *ground* the properties that distinct, co-located objects have. See Chapter 11, Section 3 of Korman’s 2015 book for details.

the arbitrariness argument against Conservatism. He maintains that there *is* an ontologically significant difference: namely, the presence of creative intentions in the case of statues and the absence of creative intentions in the case of gollyswoggles. In denying [AR33], Korman cuts off both Eliminativist and Permissivist versions of the argument.

Now, let's try out an analogous argument connecting the thermometer cases in which we are sampling with a specific stopping rule. The basic idea is to deny that there is any evidentially significant difference between the two sorts of cases. An arbitrariness argument for the claim that the ECP holds in cases like S_1 now proceeds by comparing cases like S_1 with the thermometer case:

- [ARG1] There is no epistemically significant difference between the thermometer case and cases like S_1 .
- [ARG2] If so, then: if the ECP holds with respect to the thermometer case then the ECP holds with respect to cases like S_1 .
- [ARG3] The ECP holds with respect to the thermometer case.

[ARG4] The ECP holds with respect to cases like S_1 .

The arbitrariness argument just sketched represents progress in a few ways. First, the argument gives us more than a bare intuition, which is based on a single toy example, about a rather complicated epistemological principle. Second, the argument makes explicit the basis for generalizing: that to distinguish between the toy example and real examples would be arbitrary. And third, the argument forces Wasserman (or anyone else unhappy with its conclusion) to identify a defective premiss and provide reasons for rejecting it. Specifically with respect to the third point: [ARG2] looks unassailable, and Wasserman seems to accept [ARG3]. So, he should reject [ARG1].

In order to reject [ARG1], one needs to find an epistemically significant difference between the thermometer case and cases like S_1 . Those who think intentions have ontological significance in the case of experiments are well-positioned to identify such a difference. Suppose that intentions (partially) determine which experiments exist and which properties (especially modal properties) those experiments have. Then the fact that in cases like S_1 , the experimenter *intends* to conduct a mixture experiment while in examples like the thermometer case, the experimenter *does not* intend to conduct a mixture experiment marks an ontologically significant difference between the two cases. Hence, the thermometer case is *ontologically* different from cases like S_1 . Does the ontological difference make an epistemic difference? Frequentists who have come this far are in a good position to say *yes*. Frequentists think that the modal properties of an experiment matter because they determine the sampling distribution. The thermometer case and similar thought experiments are an embarrassment because they seem to involve experiments that have modal properties that intuitively ought to be ignored but that a frequentist cannot ignore given her own inferential principles. Put another way: it seems arbitrary to treat the thermometer case differently from cases (like S_1) that have the same modal properties determining the

same sampling distribution. But the ontological difference made by creative intentions is precisely a difference in the modal properties relevant to determining what the sampling distribution looks like.

Proponents of the Likelihood Principle might charge that the reply I've just sketched is question begging. Indeed, the reply *is* question begging in the *dialectical* sense. But it is not *intrinsically* question begging.²³ That is, the frequentist asserts something denied by friends of the Likelihood Principle: namely, that the modal characteristics of cases (like S_1) involving stopping rules are epistemically significant. However, in saying that there is an epistemic difference between the thermometer case and ordinary cases involving stopping rules, the frequentist is not smuggling in the conclusion in an objectionable way, provided she can point to differences of the right sort (from her own point of view) between the modal profiles of the experiments in the two cases. By appealing to creative intentions, the frequentist can explain how the two cases differ and thus explain why it is appropriate (at least by her own lights) to treat them differently.

Return now to Gandenberger's explanation of how the thermometer case supports the ECP. He claims that when we read the thermometer case, we have an intuition that the modal profile of an experiment has no evidential significance. For Gandenberger, what justifies the ECP is "the intuition that features of experiments that could have been but were not performed are irrelevant to the evidential meaning of the outcome of the experiment that actually was performed" (482). However, if one thinks that intentions are ontologically significant with respect to which experiments exist and with respect to which properties those experiments have, then one may reasonably dissent from Gandenberger's explanation of our judgment that the coin flip is evidentially irrelevant and offer an alternative explanation: The coin flip is not evidentially relevant to the outcome of the experiment that was actually performed because *the coin flip is not part of the experiment that was actually performed in the thermometer case*. The thermometer case does not involve any experiment that is correctly described as a mixture experiment!

To be a bit more precise, what one ought to say is that the thermometer case is under-described. The thermometer case isn't explicit about what experiment you intended to perform. The natural interpretation, which justifies the claim that the coin flip is not part of the experiment actually performed in the thermometer case, is that you intended to perform an experiment in which you measured the temperature of a sample with thermometer T_2 . The coin flip was part of some external circumstances that could have thwarted your intention. But it was not part of the experiment you actually conducted. An alternative interpretation is that you *intended* to perform a mixture experiment in which a coin flip determines which of two thermometers you will use. But if so, then it is no longer surprising to find that the experiment you perform has different evidential value from the experiment John performs. By including an intention to perform a mixture experiment, we make the thermometer case ontologically—and consequently, epistemologically—similar to

²³ The distinction between dialectical and intrinsic senses of begging the question are due to Korman (2015, 28–29).

ordinary cases involving stopping rules. The two interpretations suggest the following dilemma. Either the experiment actually performed in the thermometer case is a mixture experiment (because you intended to conduct a mixture experiment) or it is not a mixture experiment (because you intended to conduct a simple experiment). If you actually performed a mixture experiment, then you ought to analyze it as you would a mixture experiment. The case is neither puzzling nor embarrassing. You should thank John for reminding you what kind of experiment you intended to perform. Alternatively, if you actually performed a simple experiment, then you ought to analyze it as you would a simple experiment. Again, the case is neither puzzling nor embarrassing. You should remind John that you didn't actually perform a mixture experiment and so shouldn't analyze what you did as if it were a mixture experiment. The case *appears* puzzling and embarrassing because of a kind of equivocation.

One might be tempted to object that regardless of your intentions in the thermometer case, you *do* all the same things. Even if you intended to perform a simple experiment, you still flipped a coin, and the coin flip has a modal profile that ought to matter to a frequentist. Going further, one might argue that intentions are screened off from having evidential import by whatever the modal facts happen to be. Once we know the modal profile of some experiment, the creative intentions that gave rise to that profile are irrelevant. If so, one might think that the thermometer case supports the ECP because the coin flip secures the relevant modal profile for the experiment. Two things ought to be conceded here: first, modal properties do screen off intentions from evidential import; and second, modal properties may exist without being brought about by any creative intentions. However, friends of creative intentions still have a response available. They may say that creative intentions matter with respect to which experiments exist and to which modal properties belong with which experiments. Such an argument might go like this. In the thermometer case, you do not intend for the modal properties of the coin to be part of the experiment. Hence, the modal properties of the coin *are not* part of the experiment. In the thermometer case, your creative intentions do not bring into existence an experiment with the relevant modal profile. They do not bring into existence a mixture experiment at all. But in cases involving stopping rules, the experimenter *does* intend to bring an experiment with the relevant modal profile into existence. The move here is analogous to what Baker and Korman say about the statue-shaped meteor. They want to say that a meteor shaped by random weathering is not a statue, even if its parts are arranged so that it is structurally identical to a statue that was produced by an artist. As Baker says, the meteor “looks like a statue [but] is not a statue.” Similarly, friends of creative intentions may say that if you did not intend for the coin flip (with its modal profile) to be part of the experiment in the thermometer case, then the experiment you perform *looks like* a mixture experiment, but it *isn't* a mixture experiment. Handling the objection in this way has the further virtue of explaining why an experimenter does not have to take into consideration her entire (chancy) history in analyzing the results of her experiments—answering a challenge Gendenberger raises against opponents of the ECP (481–482). An experimenter's history does not matter so long as that history is not intended to be part of the experiment.

3 Concluding remarks

In this paper, I have shown how one might resist two influential arguments for the Likelihood Principle by asserting the ontological significance of creative intentions. The argument from intentions, as I have formulated it, assumes that two experiments differing only in what some experimenters intended to do are identical. But that premiss should be rejected by anyone who thinks that creative intentions can matter to what there is (independently of how one arranges the furniture of the world). Similarly, axiomatic arguments in the style of Birnbaum assume a conditionality principle that is suspect if creative intentions have ontological significance.

At this point, I want to speculate a bit and suggest a possible line for future research. I think there are two plausible package deals—natural combinations of positions respecting the Likelihood Principle, abstract creationism, and the ontology of experiments—and a third that I find somewhat less plausible but that might be defensible. The first package deal maintains that experiments are eternal abstract objects (e.g. types of performance), that creative intentions do not *independently* matter with respect to what experiments there are or with respect to what experiment is performed on any given occasion, and that the Likelihood Principle is true. The second package deal maintains that experiments are abstract artifacts, that creative intentions do independently matter with respect to what experiments there are or with respect to what experiment is performed on any given occasion or both, and that the Likelihood Principle is false. The third package deal, which I find less plausible, maintains that experiments are eternal abstract objects of a peculiar sort (e.g. *indicated* types like those that appear in Levinson's account of musical works), that creative intentions do not independently matter with respect to what experiments there are but do independently matter with respect to what experiment is performed on any given occasion, and that the Likelihood Principle is false. I expect that the third package is susceptible to arguments similar to those offered by Dodd against Levinson's account of musical works and that under pressure from those arguments, the package collapses into one or the other of the first two. But I do not have well-developed, compelling arguments, yet. More work needs to be done to either verify or falsify my claim that these are the only plausible package deals and my claim that the third package deal is not really viable.

In closing, I find it encouraging that debates about abstract creationism might matter for one's philosophy of statistics and that one's opinions about the epistemology of survey sampling might matter to how one thinks about the metaphysics of artifacts. Scientific inquiry in one domain often has unexpected consequences for scientific inquiry in other domains. This is one way in which science is unified. I take the present paper to be an illustration that the same may be said of philosophy. I certainly did not expect to find that work on the ontology of artifacts, fictional characters, and musical works would be relevant to work in the philosophy of statistics. And yet, it seems that they are related. Philosophical

inquiry in one domain often has unexpected consequences for philosophical inquiry in other domains. And this is one way in which philosophy is unified.

Acknowledgements Thanks to Sam Fletcher, Konstantin Genin, Daniel Malinsky, Conor Mayo-Wilson, Greg Gandenberger, Noel Saenz, Jonah Schubach, and Kristin Seemuth-Whaley for comments on earlier drafts. Special thanks to Dan Korman for reading and criticizing multiple drafts, for helping me refine the arguments, and for motivating the project with his own excellent work.

References

- Aad, G., Abat, E., Abdallah, J., Abdelalim, A. A., Abdesselam, A., Abdinov, O., et al. (2008). The ATLAS experiment at the CERN large hadron collider. *Journal of Instrumentation*, 3, S08003.
- Anelli, G., Antchev, G., Aspell, P., Avati, V., Bagliesi, M. G., Berardi, V., et al. (2008). The totem experiment at the cern large hadron collider. *Journal of Instrumentation*, 3, S08007.
- Baker, L. R. (2007). *The metaphysics of everyday life: An essay in practical realism*. Cambridge: Cambridge University Press.
- Barrett, H. C., Laurence, S., & Margolis, E. (2008). Artifacts and original intent: A cross-cultural perspective on the design stance. *Journal of Cognition and Culture*, 8(1), 1–22.
- Berger, J. O., & Wolpert, R. L. (1988). *The likelihood principle: A review, generalizations, and statistical implications* (2nd ed.). Hayward: Institute of Mathematical Statistics.
- Birnbaum, A. (1962). On the foundations of statistical inference. *Journal of the American Statistical Association*, 57(298), 269–306.
- Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition*, 60(1), 1–29.
- Brock, Stuart. (2010). The creationist fiction: The case against creationism about fictional Characters. *Philosophical Review*, 119(3), 337–364.
- Burazin, L. (2016). Can there be an artifact theory of law? *Ratio Juris*, 29(3), 385–401.
- Caplan, B., & Matheson, C. (2004). Can a musical work be created? *The British Journal of Aesthetics*, 44(2), 113–134.
- Chaigneau, S. E., Castillo, R. D., & Martínez, L. (2008). Creators' intentions bias judgments of function independently from causal inferences. *Cognition*, 109(1), 123–132.
- Chaigneau, S. E., Puebla, G., & Canessa, E. C. (2016). Why the designer's intended function is central for proper function assignment and artifact conceptualization: Essentialist and normative accounts. *Developmental Review*, 41, 38–50.
- Deutsch, H. (1991). The creation problem. *Topoi*, 10(2), 209–225.
- Dodd, J. (2000). Musical works as eternal types. *British Journal of Aesthetics*, 40(4), 424–440.
- Dodd, J. (2002). Defending musical Platonism. *The British Journal of Aesthetics*, 42(4), 380–402.
- Dodd, J. (2004). Types, continuants, and the ontology of music. *The British Journal of Aesthetics*, 44(4), 342–360.
- Dodd, J. (2007). *Works of music: An essay in ontology*. Oxford: Oxford University Press on Demand.
- Edwards, A. (1972). *Likelihood*. Cambridge: Cambridge University Press.
- Friedell, D. (2016). Abstract creationism and authorial intention. *The Journal of Aesthetics and Art Criticism*, 74(2), 129–137.
- Gandenberger, G. (2015). A new proof of the likelihood principle. *The British Journal for the Philosophy of Science*, 66(3), 475–503.
- Hilpinen, R. (1993). Authors and artifacts. *Proceedings of the Aristotelian Society*, 93, 155–178.
- Hilpinen, R. (2011). Artifact. In Edward N. Zalta (Ed.) *The Stanford encyclopedia of philosophy* (Winter 2011 Edition). <http://plato.stanford.edu/archives/win2011/entries/artifact/>.
- Houkes, W., & Vermaas, P. E. (2009). Contemporary engineering and the metaphysics of artefacts: Beyond the artisan model. *The Monist*, 92(3), 403–419.
- Howson, C., & Urbach, P. (1996). *Scientific reasoning: The Bayesian approach*. Chicago, IL: Open Court Publishing.
- Irmak, N. (2013). Software is an abstract artifact. *Grazer Philosophische Studien*, 86(1), 55–72.
- Korman, D. Z. (2014). The vagueness argument against abstract artifacts. *Philosophical Studies*, 167(1), 57.
- Korman, D. Z. (2015). *Objects: Nothing out of the ordinary*. Oxford: Oxford University Press.

- Korman, D. Z., & Carmichael, C. (2017). What do the folk think about composition and does it matter? In D. Rose (Ed.), *Experimental metaphysics*. London: Bloomsbury Academic.
- Kornblith, H. (1980). Referring to artifacts. *The Philosophical Review*, 89(1), 109–114.
- Levinson, J. (1980). What a musical work is. *The Journal of Philosophy*, 77(1), 5–28.
- Malt, B. C., & Sloman, S. A. (2007). Artifact categorization: The good, the bad, and the ugly. In E. Margolis & S. Laurence (Eds.), *Creations of the mind: Theories of artifacts and their representation* (pp. 85–123). Oxford: Oxford University Press.
- Mayo, D. (1996). *Error and the growth of experimental knowledge*. Chicago, IL: The University of Chicago Press.
- Mayo, D. G., & Kruse, M. (2001). Principles of inference and their consequences. *Foundations of Bayesianism* (pp. 381–403). Netherlands: Springer.
- Pollard, S. (2007). Mathematical determinacy and the transferability of aboutness. *Synthese*, 159(1), 83–98.
- Rose, D., & Schaffer, J. (2015). Folk mereology is teleological. *Noûs*, 51, 238–270.
- Rouder, J. N. (2014). Optional stopping: No problem for Bayesians. *Psychonomic Bulletin and Review*, 21(2), 301–308.
- Saenz, N. B. (2015). A grounding solution to the grounding problem. *Philosophical Studies*, 172(8), 2193–2214.
- Sainsbury, M., & Tye, M. (2012). *Seven puzzles of thought and how to solve them*. Oxford: Oxford University Press.
- Salmon, N. (1998). Nonexistence. *Noûs*, 32(3), 277–319.
- Savage, L. (1962). *The foundations of statistical inference*. New York, NY: Wiley.
- Schwartz, S. (1978). Putnam on artifacts. *The Philosophical Review*, 87(4), 566–574.
- Soames, S. (2002). *Beyond rigidity: The unfinished semantic agenda of naming and necessity*. Oxford: Oxford University Press.
- Sorge, R. E., Martin, L. J., Isbester, K. A., Sotocinal, S. G., Rosen, S., Tuttle, A. H., et al. (2014). Olfactory exposure to males, including men, causes stress and related analgesia in rodents. *Nature Methods*, 11(6), 629–632.
- Sprenger, J. (2009). Evidence and experimental design in sequential trials. *Philosophy of Science*, 76(5), 637–649.
- Steele, K. (2013). Persistent experimenters, stopping rules, and statistical inference. *Erkenntnis*, 78(4), 937–961.
- Thomasson, A. L. (1999). *Fiction and metaphysics*. New York, NY: Cambridge University Press.
- Thomasson, A. L. (2003). Realism and human kinds. *Philosophy and Phenomenological Research*, 67(3), 580–609.
- Thomasson, A. L. (2007). Artifacts and human concepts. In E. Margolis & S. Laurence (Eds.), *Creations of the mind: Theories of artifacts and their representation* (pp. 52–73). Oxford: Oxford University Press.
- Thomson, J. (2008). People and their bodies. In T. Sider, J. Hawthorne, & D. W. Zimmerman (Eds.), *Contemporary debates in metaphysics*. London: Blackwell.
- Van Inwagen, P. (1977). Creatures of fiction. *American Philosophical Quarterly*, 14(4), 299–308.
- Van Inwagen, P. (1990). *Material beings*. Ithaca, NY: Cornell University Press.
- Vecsey, Z. (2014). Linguistic data and fictional characters. *Linguistic and Philosophical Investigations*, 13, 140.
- Wasserman, L. (2012). Statistical principles? *Normal Deviate* (Blog). normaldeviate.wordpress.com/2012/07/28/statistical-principles.
- Yu, E. C., Sprenger, A. M., Thomas, R. P., & Dougherty, M. R. (2014). When decision heuristics and science collide. *Psychonomic Bulletin and Review*, 21(2), 268–282.
- Zvolenszky, Z. (2012). Against Sainsbury's Irrealism About Fictional Characters: Harry Potter as an Abstract Artifact. *Hungarian Philosophical Review (Magyar Filozófiai Szemle)*, 4, 83–109.