Is Thermodynamics Subjective?

Katie Robertson¹ and Carina Prunkl²

¹Department of Philosophy, University of Stirling, Stirling, UK and ²Department of Philosophy, University of Oxford, Oxford, UK

Corresponding author: Katie Robertson; Email: katie.robertson@stir.ac.uk

(Received 13 February 2023; revised 05 March 2023; accepted 08 March 2023; first published online 29 March 2023)

Abstract

Thermodynamics is an unusual theory. Prominent figures, including J. C. Maxwell and E. T. Jaynes, have suggested that thermodynamics is anthropocentric, and contemporary approaches label thermodynamics a “subjective theory.” Here, we evaluate the arguments for anthropocentrism but conclude that instead of pointing to an anthropocentric view, they point towards a resource-relative understanding of thermodynamics which can be shorn of the “subjective gloss.”

1. Introduction

If physical theories were people, thermodynamics would be the village witch. The other theories find her somewhat odd, somehow different in nature from the rest, yet everyone comes to her for advice, and no one dares to contradict her.

(Goold et al. 2016, 1)

One respect in which the village witch differs from other theories is that many question her objectivity. Jaynes (1957), for example, argues for the “anthropomorphic” nature of entropy. Lloyd (2006) writes that “entropy is the information we don’t have, and is therefore subjective.” Bridgman (1941, 214) remarks that “thermodynamics smells of its human origins.”

At the same time, many hold thermodynamics in high regard. Eddington famously wrote that “if your [pet theory of the universe] is found to be against the second law of thermodynamics I can give you no hope; there is nothing for it but to collapse in deepest humiliation” (Eddington 1935, 81). If thermodynamics is not objective, some draw drastic consequences: “Such a view would create some profound philosophical problems and would tend to undermine the objectivity of the scientific enterprise” (Denbigh and Denbigh 1985, vii). In this article, we argue that these challenges to the objectivity of thermodynamics can be overcome.
We begin, in section 2, with a discussion of “objectivity,” and argue that the key question is whether thermodynamics is anthropocentric in a way different from other physical theories. In section 3, we argue that—whilst a compelling view at the time Maxwell was writing—developments in quantum statistical mechanics reveal that there are ways to avoid the anthropocentrism Maxwell endorsed. Nonetheless, his insights continue to illuminate the frontier of quantum thermodynamics. This resource-relative approach helps reveal important insights about thermodynamics—we discuss this in section 4. Yet, the similarities between this view of thermodynamics and special relativity leave it open to attack as “subjective” by the lights of Nozick (2001). In section 5 we argue this is mistaken: we conclude that thermodynamics is non-fundamental but nonetheless objective.

2. Objectivity: What is at stake?

Is thermodynamics objective? This depends on what we mean by “objective”—there are many options (John 2021). Some views can be set aside here; no one is concerned that non-cognitive values prevent thermodynamics from being objective. Instead, a different general theme is relevant: objectivity as “faithfulness to facts” (Reiss and Sprenger 2020) or as reality independent of the existence of the human minds. The failure of some feature to be objective brands it as “subjective.” In this section, we set aside the most alarmist worry about thermodynamics, and discuss how bringing in “agents” might make thermodynamics “subjective.”

Denbigh and Denbigh suggest that should thermodynamics be found not to be objective, this would threaten to “undermine the scientific enterprise.” They say this because objectivity is meant to be something good about the scientific enterprise; scientific knowledge is held in high esteem because it is objective, thus trustworthy, robust, and impartial.

But, in reality, no one is worried that thermodynamics or statistical mechanics are not objective in the objectivity-as-robustness or trustworthiness or impartial sense. The standing of thermodynamics as a successful scientific theory is undisputed; the process through which we discovered thermodynamics is not somehow worse than the rest of physics.

Some discussions of thermodynamics seemingly threaten objectivity by bringing in epistemic considerations; in particular there is often talk of “agents” or “observers.” If the differences between certain agents are relevant to thermodynamics, thermodynamics might be subjective in the sense of depending on the particularities of certain agents, such as their credences. For example, if different agents assign different probability distributions, then variations between agents may make thermodynamics “subjective” in the sense of depending on particular subjects.

Yet, it seems this variation between agents is not what most critics have in mind. David Albert rhetorically questions “why would our ignorance bring it about that milk dissolves in coffee?” Albert’s criticism is not that my (rather than your) ignorance or beliefs is the problem; the issue is not intersubjective variation of ignorance but the introduction of ignorance at all.

The concern is that by discussing agents (or observers), the theory is not mind-independent: without them, thermodynamics would not be an enlightening
description of the world. But moreover, the agents might need to be agents like us; recall that Bridgman was concerned about the “human smell” of thermodynamics, and Denbigh and Denbigh were concerned about “human ignorance.” The worry is that thermodynamics is anthropocentric.

Merely mentioning “observers” or “agents” does not bring anthropocentrism knocking. Special relativity discusses observers (a synonym for “reference frame”); quantum mechanics discusses “observables” and mentions observers in textbook expositions of the famously problematic “spontaneous collapse upon measurement.” Decoherence has largely removed the need to “insert the human perspective deep within physics.” In special relativity, likewise, “observers” have little to do with us; an observer just indexes a reference frame, i.e., a geometrical convention of how spacetime is split into space and time.

The use of “observers” in special relativity and quantum mechanics needn't bring any concern about anthropocentrism. But many claim thermodynamics does. Before we evaluate that claim, what is it for a feature to be subjective in the sense of anthropocentric? Nozick’s (2001) account takes the root of objectivity to be invariance across perspectives; if a quantity is invariant when transforming from one perspective to another, then it is more objective than a quantity that varies. Is a particular feature invariant across different types of creature? The visual appearance of flowers varies between humans and bumblebee; it is not invariant, hence not objective in this sense. If a feature of the world appears to humans in a certain way, but to other creatures in a different way, that feature is anthropocentric.

How concerning this is depends on your stance on scientific realism; if thermodynamics is anthropocentric this precludes standard scientific realism about thermodynamics. Conversely, if thermodynamics is unavoidably anthropocentric, then the fact that such an important physical theory has this feature bolsters the case for pragmatism more generally.

3. Maxwell’s charge of anthropocentrism

If thermodynamics is anthropocentric, some features will differ with different types of creatures. One type of putative creature is the famous Maxwellian demon, from the perspective of which the second law is not a genuine regularity. Is such a creature possible? Maxwell thought so. We discuss—and reject—this view in section 3.2. But first, in section 3.1 we consider the heat/work distinction that is so central to thermodynamics, yet one that Maxwell claimed is thoroughly anthropocentric.

3.1. Maxwell part 1: Heat and work

The heat/work distinction—or least a distinction between different types of energy transfer—is central to phenomenological thermodynamics. Maxwell describes the underpinning of heat/work distinction in thoroughly anthropocentric terms:

---

1 “Agents” needn’t immediately lead to subjectivity; cf. Reutlinger (2021), who cashes out this invariance under the biases of the individual scientist. But even if the attributes of particular scientists do matter, provided that a range of attributes are considered, objectivity is possible, following Longino (1990).

2 Caratheordory’s axiomization of thermodynamics removes “heat” as a primitive term; Uffink argues that nonetheless some kind of distinction about energy transfer is required.
“Dissipated energy is energy we cannot lay hold of and direct at pleasure, such as the energy of the confused agitation of molecules which we call heat. Now, confusion, like the correlative term order, is not a property of material things in themselves, but only in relation to the mind which perceives them” (Maxwell [1878] 1965, 220).

The claim that “confused” motion (heat) and “ordered” motion (work) depends on particular minds makes the heat/work distinction anthropocentric. Maxwell ties it to our particular perceptual and cognitive perspective: “It is only to a being in the intermediate stage, who can lay hold of some forms of energy while others elude his grasp, that energy appears to be passing inevitably from the available to the dissipated state” (ibid). For another type of being, both forms of energy may be in their grasp—i.e., within their capabilities, as will be discussed in the next section—and the difference between the types of energy might not be apparent.

From the perspective of the underlying theory both heat and work correspond to (a form) of molecular motion. Whether that motion counts as ordered (“work”) or disordered (“heat”) depends on what we consider to be ordered or not. A much smaller creature might not draw the distinction in the way we do—according to this charge of anthropocentrism.

Merely looking at the microscopic level of molecules jiggling around indeed suggests that both heat and work are “nought but molecules in motion” (Maxwell [1874] 1884). But as Uffink emphasizes, there are more “lower-level” resources available than the bare description of the motion of molecules; “the distinction between heat and work can be framed in a mind-independent way with the help of concepts from probability theory” (Uffink 1996, 344).

Probabilities—as are inevitable in statistical mechanics (SM)—bring in their own whiff of the human smell; largely because probabilities in SM are commonly associated with Jaynes, who took a thoroughly subjectivist view of them. Yet, popular though Jayne’s view is, there are more objectivist views available (cf. Myrvold 2021).

 Nonetheless, to avoid well-known problems with probabilities in classical SM, we focus on the quantum case here. Probabilities are an inherent part of quantum mechanics, and so are not an addition to the formalism. For more on the claim that SM probabilities are identical to or emerge from quantum mechanics probabilities see Popescu et al. (2005).

To derive the first law of thermodynamics—and hence the distinction between heat and work—from quantum mechanics, we will assume that it is acceptable to identify thermodynamic entities (such as heat, work, entropy, etc.) with expectation values. The (mean) energy of a quantum state $\rho$ is given by the expectation value of its Hamiltonian, $\langle H \rangle_\rho = \text{Tr}[H\rho]$. We now consider a quantum system $\rho_S$ that interacts with an environment $\rho_E$. The joint system initially is a product state, i.e., $\rho(t) = \rho_S(t) \otimes \rho_E(t)$, and evolves under the Hamiltonian $H = H_S + H_E + V_{SE}$. Pending further assumptions on the initial conditions and the Hamiltonian, we then derive the following equation for the system’s energy change (Maroney 2007):

---

3 In classical SM, the use of expectation values, unlike Boltzmannian approaches, is embraced by Gibbsian approaches (cf. Myrvold forthcoming) as this incorporates fluctuations. In the quantum case this dispute is fairly moot; expectation values are unavoidable.
where the term $Q_{HS} = \int_{t_0}^{t_1} \langle [H_s, V_{SE}] \rangle \rho(\sigma) dt$ describes the interaction-induced energy increase in the system and notably depends on the interaction potential $V_{SE} \in \mathcal{O}(\mathcal{H}_S \otimes \mathcal{H}_E)$. Equation (1) can be integrated and rewritten as the familiar first law, i.e., $\Delta U_S = W_E + Q$. In the way presented above, $W_E$ and $Q$ can be considered to represent the quantum-mechanical analogues of work and heat. To add to the idea that these quantum expressions do play the heat and work role, note that there is an upper limit on how much energy can be extracted on average as work from the system.\(^4\) This is what is known as adiabatic accessibility: “not all of the energy of a system is available for work” (Maroney 2007, 22).

Thus, we can understand heat and work at a more fundamental level without appealing to distinctions such as “order/disordered.” Yet, the presence of expectation values means that quantum probabilities are involved. Does this raise the spectre of anthropocentrism? Whilst the nature of quantum probabilities is interpretation dependent, most agree that quantum probabilities are the most objective probabilities in nature. Thus, the underlying basis of heat/work distinction needn’t be anthropocentric.

3.2. Maxwell part 2: Demons and the second law

Even if the heat/work distinction is not anthropocentric, there is another threat: Maxwell’s demon. This demon is usually viewed as a challenge for the validity of the second law; why is it relevant to whether thermodynamics is anthropocentric? The connection comes from the sense in which the demon is possible or not, which motivates a wider part of Maxwell’s view that thermodynamics is linked to our capabilities.

There are a variety of different types of possibility. Whether something is physically possible depends on which theory you adjudicate it with respect to. According to Newtonian physics it is possible for an object—such as my car—to travel faster than light. But, of course, according to special relativity it is not possible. In addition to the possibilities indexed to theories, we can also consider practical possibilities: it is not practically possible for my old car to travel at 200 miles per hour, though it is physically possible according to special relativity.

In what sense is Maxwell’s demon possible, if at all? From the perspective of an underlying time-reversal invariant theory, returning a system to its earlier (and lower entropy) state looks possible. Imagine flipping the momenta of each of the gas molecules, as Loschmidt suggested; a Maxwellian demon appears possible according to classical or quantum mechanics. But there is an important caveat: a Maxwellian demon only truly violates the second law if there is a decrease in thermodynamic entropy for the entire (isolated) system. After all, the thermodynamic entropy of a subsystem decreases in an isothermal compression, but there is a compensating increase in the entropy of the heat bath. So the question is: does the operation of the Maxwellian demon lead to an increase in thermodynamic entropy elsewhere?

\(^4\) A further condition: the Hamiltonian must be bounded from below.
The answer to this question is fiercely disputed. We will discuss the principle behind this controversy—Landauer’s principle (LP)—but first we will examine Maxwell’s own views.

Maxwell believed that there was no physical principle (as LP claims to be) preventing us from acting like the demon does. That is, there is no underlying law (in SM or other physical theories) ruling out the possibility of a Maxwellian demon. As Myrvold writes, “For Maxwell … it is only our current, but perhaps temporary, inability to manipulate molecules individually that prevents us from doing what the demon would be able to do” (Myrvold 2011, 238). It is a mere matter of practical (im)possibility that we are too clumsy to act like Maxwell’s demon. Maroney (2009) agrees; much like Boltzmann thought it would be practically impossible to reverse the momenta of the molecules in a gas like Loschmidt suggested, it is impractical for us to act like the demon.

But the horizons of the practically possible constantly shift as technology progresses; in Maxwell’s times crossing the Atlantic in a matter of hours was not possible, yet today it is. For Loschmidt’s demon, we do now have the abilities to reverse momenta (or something closely analogous), at least in the spin echo experiment (Hahn 1950). What is practically possible changes over time; Maxwell imagined that in (his) future we could perhaps manipulate individual molecules, and thus successfully violate the second law. The second law is not invariant under the different abilities of agents; thermodynamics once again is subjective. Arguably, this is a stronger subjectivity than just depending on human capabilities, since it is not just limited to our innate size and perspective but rather on our technology, which changes over time. Consequently, there is less invariance and so less objectivity.

But, unlike in Maxwell’s day, we can manipulate individual molecules. Does this mean that we can violate the second law? There are certainly a range of devices both engineered and in nature that are claimed to act like the demon (Serreli et al. 2007). But these are mere illusionists rather than true demons—if there is a cost to their operation as Landauer’s principle decrees.

Landauer’s principle claims that there is an entropic cost of $k_B \ln 2$ to resetting one bit of data. Why is this relevant to Maxwell’s demon? Earlier, we emphasized the importance of the system–environment split: the demon only violates the second law if the total entropy of the system, environment, and demon decreases.

Yet Landauer’s principle has a strange status; it is gospel within parts of the physics community (and motivates bringing information-theoretic considerations in thermodynamics; an industry that is now its own subfield, cf. del Rio et al. 2015). On the other hand, Landauer’s principle is heresy—according to some philosophers of physics (Norton 2011). There is a heated debate: some proofs have been justly criticized by Earman and Norton, since they do not establish the principle in full generality as they rely on very specific mechanisms. (Although see Myrvold forthcoming).

Should we believe Landauer’s principle to be true? In one sense, if Landauer’s principle is false and a Maxwellian demon is possible, this is fabulous news: we can solve the energy crisis! (Cf. Wallace 2014.) The second law would be dethroned as one of the great principles of science. But even if there are ingenious mechanical devices that appear to operate like the demon, no one thinks these hold the key to solving the energy crisis.
Greater-than-Carnot efficiency engines are but a dream. If anyone thought that meaningful violations of the second law were possible, this would be a hive of research.\footnote{We add “meaningful” violations; whether very small violations are possible depends on one’s construal of the second law.}

If we take Landauer’s principle—or other exorcisms (Norton 2014)—to vanquish the demon, where does that leave the question of objectivity? There are good reasons to rule out a Maxwell demon that allows for greater-than-Carnot efficiency engines. As such, the second law doesn’t depend on our inability to manipulate molecules; the fact we now can doesn’t change the maximum efficiency of engines.

4. Maxwell’s legacy: Means-relative, resource theories, and control theory

We have seen that neither the heat/work distinction nor the specter of a Maxwellian demon require us to view thermodynamics as anthropocentric. Yet, Maxwell has left an important legacy by framing thermodynamics in terms of what operations one can or cannot perform on a thermodynamic system.

We begin by introducing three prominent approaches to thermodynamics: means-relative, control-theoretic, and resource-theoretic. Given the similarities between them, we will jointly address them as resource-relative.

Resource theory: Resource theories have emerged within recent years from within the field of quantum thermal physics. They focus on what (quantum) states can be reached via a fixed set of operations, and are based on three main ingredients (Brandão and Gour 2015): (a) a set of “free states” that are abundant and freely available to the experimenter, (b) a set of available quantum operations, and finally (c) a set of “resources”: states that cannot be created by (a) and (b) alone. What makes resource theories so impressive is that for some intuitive choices of (a), (b), and (c), the quantum mechanical analogue of the second law of thermodynamics can be derived (Brandão et al. 2013).

Means-relative thermodynamics: Myrvold (2011) characterizes Maxwell’s approach as means-relative: heat, work, and entropy are defined with respect to the means an agent has to manipulate the system. In the same breath, Myrvold emphasizes that while means can vary between agents, “it would be misleading to call them subjective, as we are considering limitations on the physical means that are at the agent’s disposal. On Maxwell’s view, the distinction between work and heat is means-relative” (Myrvold 2011, 239).

Control theory: Control theory origins from engineering and deals with the control of dynamical systems and automatic feedback loops to achieve automatic control. Interestingly, Maxwell is often listed as a pioneer of control theory (Dorf and Bishop 2018, 37). In 1868, Maxwell formulated a mathematical theory related to “governors”—devices to control the speed of engineering systems, e.g., engines (Maxwell 1868). According to Wallace (2014), who argues for a control-theoretic approach to thermodynamics, a control theory is a “theory of which transitions between states can be induced on a system (assumed to obey some known underlying dynamics) by means of operations from a fixed list” (15).

Common to all three theories is that they start out with a fixed number of operations that an experimenter can or cannot perform on a system. In the
means-relative case, these operations are implicitly determined by the abilities of the agent. The other two theories explicitly list the operations available to the agent.

Resource-relative approaches to thermodynamics emphasize the importance of external interventions on thermodynamic systems. Such interventions are needed since the state space of thermodynamics is one of equilibrium states where—by definition—macroparameters are unchanging. Thus, for a system’s state to change at all, an intervention is required (e.g., the pushing of a piston to isothermally compress a gas or the putting in contact with a heat bath). By depending on external interventions for dynamics, thermodynamics differs from other physical theories, such as Newtonian mechanics or general relativity.

The focus on external interventions invites talk about agents, yet such agents by no means need to be human. As Myrvold (2020) points out, “manipulability needn’t be manipulable by us” (7). The focus on interventions alone, therefore, does not require us to consider thermodynamics a subjective theory. Yet, consideration of the Gibbs paradox, and the resource-relative resolution of it (Gibbs 1878; Jaynes 1965), leads to the view that entropy looks subjective in the following way. Alice has the resources to distinguish between the two gases separated by a partition, whereas her friend Bob does not. For Alice, there is an entropy of mixing $\Delta S = nR\ln 2$ since she has lost the ability to do work (by inserting semi-permeable membranes), but for Bob there is no change in entropy.

This resource-relative view understands entropy as restricting the maximum amount of work that can be extracted from the system, given a set of operations that act on the thermodynamic state. Faist and Renner (2018, 11) have furthermore shown that it is possible to “switch” between different resource frames, or “different observers,” by using completely positive, trace-preserving maps to recover one from the other. With the help of a “recovery map,” Bob’s picture can be recovered from Alice’s picture.

In light of the above, it becomes clear why one might be tempted to take thermodynamics as subjective. If the very state of a system depends on the resources we have available, how could thermodynamics possibly be objective? Some authors make the explicit connection: Faist and Renner (2018, 2) state that their framework “naturally treats thermodynamics as a subjective theory, where a system can be described from the viewpoint of different observers.” The pull towards subjectivity seems strong.

Next, we argue that this pull can be resisted; we argue to reject the view of subjectivity upon which it rests.

5. Against relativity-as-subjectivity

The resource-relative view emphasizes the features thermodynamics is relative to; indeed, Faist and Renner explicitly cast thermodynamics in the mold of special relativity. But is the resource-relative view in any way subjective? Faist and Renner (2018) claim so. And this immediately connects to Nozick’s account of objectivity (Nozick 2001).

Earlier, we glossed Nozick’s view as “objectivity is invariance under transforming from one perspective to another.” This is a liberalized version of Nozick’s view. Nozick originally requires invariance under admissible transformations; his archetypal example is the Lorentz transformations. But, here, we follow Stephen John’s interpretation on which talk of Lorentz transformations is a “metaphor for establishing a notion of objective claims as claims whose truth or acceptability

https://doi.org/10.1017/psa.2023.58 Published online by Cambridge University Press
does not vary with the perspective of particular (communities of) inquirers,” which John notes has striking (and perhaps surprising) resonances with Sandra Harding’s view of invariance across standpoints.

But to return to the strict reading of Nozick, since the Lorentz transformations are the exemplar admissible transformations, Lorentz-invariant quantities (e.g., spatiotemporal intervals) are more objective than quantities that vary (e.g., temporal intervals). Relative quantities are less objective than absolute quantities. Thus, if some thermodynamic quantities are “resource-frame dependent” as suggested in the previous section, then by Nozick’s lights they are less objective. Even if thermodynamics is not anthropocentric, it might still have such “metaphysical subjectivity.”

However, this literal reading of Nozick is undesirable for reasons independent of thermodynamics: it returns counter-intuitive, or unhelpful, verdicts even in its paradigm case of special relativity.

Temporal intervals between spacetime events are frame-dependent, but once we specify which reference frame the temporal interval is being specified with respect to, it is no longer frame-dependent. All observers agree that in frame X, the temporal interval is $\Delta t_X$. Indeed, the same information is conveyed by (i) specifying the spatiotemporal interval and (ii) specifying the temporal and spatial intervals with respect to a certain frame. So what is the difference between (i) and (ii)? (ii) is less mathematically elegant for sure, and less useful for building future theories. But should we say that (ii) is less objective than (i)? Given that they revealed the same facts about the world, it is hard to see how (ii) is any “less faithful to the facts”—especially when we remember that reference frames in special relativity are just geometric conventions splitting spacetime into space and time.

Instead of saying (ii) is less objective than (i), we should say it is less fundamental. For the resource-relative view of thermodynamics, this means that insofar as the analogy with special relativity holds, certain quantities are relative. Yet, relative/absolute needn’t line up with subjective/objective, and so this should not give us reason enough to think thermodynamics is subjective.

But is objectivity closely aligned with fundamentality? According to some, such as Sider, yes. This view resonates with Nozick’s view. Yet note that no one suggests that thermodynamics is fundamental (at least in one sense of the word)—it is an archetypal special science, and so conflating fundamentality with objectivity is unhelpful for the debate at hand here.

Moreover, objectivity-as-fundamentality is strikingly at odds with the purposes for which most care about objectivity: the status of the scientific enterprise as (aiming to be) epistemically virtuous, and deserving of our trust. Objectivity is closely tied to epistemic risk, and from this angle, thermodynamics is thoroughly objective, with very little epistemic risk in sight. For instance, Einstein, as quoted in Klein (1967), said: “[Thermodynamics is] the only physical theory of universal content concerning which I am convinced that, within the framework of the applicability of its basic concepts, it will never be overthrown.” Thermodynamics might not be fundamental, but it is objective.

References