January 19, 24, 2025.

Too Good to be True: Entropy, Information, Computation and the Uninformed Thermodynamics of Erasure

John D. Norton
Department of History and Philosophy of Science
University of Pittsburgh

Word count, excluding references: 4,202

The notion that theories of information and computation can augment and even complete thermodynamics has proven too enticing for many to resist, even though careful analysis has long shown that the notion fails. In so far as the results of this information-computation theoretic literature succeed, they are merely tendentious relabeling of mundane thermodynamics. When they go beyond it, they fail. The difficulties include an unsustainable conflation by Landauer's principle of the dynamic probabilities of thermalization with the static probabilities of memory devices. The most serious failure is an enduring neglect of the import of thermodynamic fluctuations.

1. Introduction

What if an excursion into the abstract realms of information and computation could extend and perhaps even complete thermodynamics, the physical theory of heat? Leon Brillouin (1961, p. 318) argued that his "negentropy principle of information is a generalization of Carnot's principle [the second law of thermodynamics]." Charles Bennett (1987, p. 108) exulted:

> The correct answer—the real reason Maxwell's demon cannot violate the second law—has been uncovered only recently. It is the unexpected result of a very different line of research: research on the energy requirements of computers.

The unexpected result is Landauer's principle, whose application to the demon is summarized as (p. 116):

> … in order to observe a molecule, it must first forget the results of previous observations. Forgetting results, or discarding information, is thermodynamically costly.

In a scientific era replete with discoveries of the extraordinary, these are ideas too intriguing and too enticing for many to resist. If Einstein can equate mass and energy, gravity and spacetime geometry, and quantum theory can allow consciousness to collapse a quantum state, why not merge the science that governs heat engines and power plants with abstract theories of information and computation? For almost a century, this idea has been celebrated in an accelerating literature that continues to grow in volume and variety.[1] Its results are extraordinary. We are assured by Jerome Rothstein (1951, p. 173), for example, that "…all physical laws become relationships between types information …"

It is a hard-won, life lesson that not every enticing idea is correct. If it looks too good to be true, it likely is not, as is the case here. This paper will summarize, necessarily briefly, reasons for a more sober appraisal of this overall literature. It has, I will argue, allowed an enthusiasm for an enticing idea to eclipse much-needed skepticism.

Section 2 below will review early instances of the core foundational difficulty that persists, even in the latest literature: its results are sustainable only when they assign tendentious labeling to otherwise mundane results in statistical physics. Claims that assert more prove unsustainable. Section 3 will recount what I believe is the most important, technical failing of this literature, its enduring neglect of thermal fluctuations. Section 4 will review how Landauer's principle tries unsuccessfully to attach an entropy cost to the erasure of information by conflating the dynamic probabilities of thermalization with the static probabilities of memory devices.

---

[1] See Leff and Rex (2003) for an early, expansive survey.

2. Entropy's New Clothes

That this literature exceeds in its ambitions has long been recognized by a persistent undercurrent of dissent. The immediate complaint is the obvious one that the realm of heat and the realm of ideas of information and computation are distinct and that claims of non-trivial connections between them must depend on equivocation. Ter Haar (1966, p. 224, his emphasis) warned: "It must be stressed here that the entropy introduced in information theory is *not* a thermodynamical quantity and that the use of the same term is rather misleading." Carnap (1977, p. 73) noted that Brillouin's treatment of entropy "makes [entropy] a logical rather than a physical concept."

Earman and Norton (1998, 1999) sought to sharpen this hesitation with a dilemma for the literature's analysis of Maxwell's demon (1998, p. 436, emphasis in original):

> In so far as information theory can protect the Second Law of thermodynamics from Maxwell's Demon by *sound* argumentation, it does so through the presumption that the Second Law must govern a naturalised Maxwell Demon. Thus the *sound* exorcism adds nothing of fundamental principle to the Second Law. It is at best a picturesque way to tease out some of its consequences. In so far as information theory provides a *profound* exorcism, it must do so by invoking hitherto neglected and novel physical principles. But what the exorcism literature has failed to present and, we believe, cannot present are compelling, independent reasons for accepting the new physical principles that connect information and thermodynamic entropy.

My overall impression is that the careful literature almost invariably accepts the *sound* horn; and correctly so. This includes analyses that may not explicitly call upon the second law, but adopt standard results in thermal physics that are equivalent to it.

This universal acceptance of the *sound* horn is tantamount to abandoning the pretense that this literature has novel results of foundational importance. It concedes that its results work backwards to determine the entropy creation that must occur in various processes on the assumption of the prior thermodynamic theory. Of course, a Maxwell's demon must be doing something dissipative, since otherwise the second law of thermodynamics would be breached, which is *by assumption* not the case. This is so, no matter how the demon is designed, whether its operation can be anthropomorphized as an observer collecting information (Szilard); or a computer forgetting information (Bennett); or as a ratchet and pawl accumulating fluctuations

(Feynman); or as a trapdoor passively responding to molecular collisions (Smoluchowski). This literature has found no single novel principle that can accommodate the variety of these cases; and none is needed once the *sound* horn is accepted.

Kenneth Denbigh's (1981, p. 113, his emphasis) assessment captures the transparency of entropy's new clothes: "… [Brillouin's notion of] 'bound information' is really nothing more than a *name* given by Brillouin to [an] entropy change…" Landauer's (1991) associated and celebrated claim in the title of his paper, "Information is Physical," appears to me to be a brash attempt to pretend that a mere renaming is somehow a major physical discovery.

Charles Bennett (2003, p. 508-509), in a response to the dilemma, accepts the *sound* horn, but counters that Landauer's principle of the thermodynamics of computation "serves an important pedagogic purpose." It saves students from a misconception that afflicted "giants like von Neumann, Gabor, and Brillouin and even, perhaps, Szilard." It is "the informal belief that there is an intrinsic cost of order kT for every elementary act of information processing."

Bennett and Landauer's pedagogy is dubious. As we shall see in the next section, Bennett and Landauer's efforts to establish that this is indeed a misconception are themselves fatally flawed by a neglect of fluctuations. The whole episode indicates a shift of foundational commitments, driven by little more than an overreaching attempt to promote computational conceptions. There are no novel experiments driving the change. We are to suppose that the giants—von Neumann!—were just confused or negligent and that a little more thought completely overturns their insights. My note here reviews how precisely the same problem afflicts Bennett and Landauer's analysis. This is not a literature with stable foundations.

As for the motivating problem of Maxwell's demon, this literature has produced an enduring debate over whether information and computational notions enable the demon's exorcism. The effect has been malign. It has distracted us, and here I ruefully include myself, from the direct question of whether a Maxwell's demon is physically possible. As a result, we long overlooked a short, simple argument (in Norton, 2018) that demonstrates the demon's impossibility using the Liouville theorem or its quantum analog. It is quite general since it does not require the highly restrictive assumption that a Maxwell's demon can be conceived as an information processor or one that must erase a memory device. In so far as the information-computation literature can licitly preclude a Maxwell's demon, it will do it by mere relabeling of the simple result just stated. In so far as it goes beyond, it errs.

3. The Neglect of Fluctuations

An enduring and serious oversight of this entire tradition is its mishandling of thermal fluctuations, also known as noise or, in radio, static. Their careful treatment gives a more precise result that vindicates the looser claim, mistakenly denounced by Bennett: "there is an intrinsic cost of order *kT* for every elementary act of information processing."

Marian Smoluchowski (1912) provided a serviceable exorcism of Maxwell's demon. He argued through numerous, still familiar examples that thermal fluctuations would fatally disrupt the mechanisms of any Maxwell's demon. This same exorcism works for the naturalized demon of Szilard's (1929) one-molecule gas engine. In work that initiates the modern tradition of information-theoretic analysis, Szilard argued that the demon is defeated by the entropy costs arising in the demon's measurement of the molecules' position. However, no novel principle concerning measurement is needed. This cost can be subsumed under the more general result that all such molecular-scale operations are entropically costly, since they must suppress fluctuations. Szilard gave a brief and frivolous strategy for dismissing the effects of thermal fluctuations by supposing (p. 304) that his one-molecule gas acts on a massive, fast-moving piston. Norton (2013, §5) explains why this fails.

Leon Brillouin also argued that the thermodynamic cost of measurement would defeat the operation of a Maxwell's demon. Yet his most celebrated example, Brillouin's "torch" (1951, 334), does not show it. The exorcism depends on computing the high thermodynamic cost of introducing a photon of sufficient energy that it can be seen above the background of thermal radiation and be used by the demon to detect the position of a molecule. This cost is not calculated using information-theoretic considerations. It is computed by requiring that creation of the photon is an entropy-creating process, sufficiently dissipative to suppress the disruptive effects of the thermal fluctuations of the background black body radiation field.

The neglect of fluctuations by the Landauer-Bennett computational tradition is far more serious since it directly undermines their most basic proposal. It is the idea that logically reversible computations can be implemented by thermodynamically reversible processes, that is, processes that are least dissipative; and this is so no matter how many steps are employed to complete the process.

What has been overlooked, repeatedly, is that thermal fluctuations preclude the completion of *any* molecular-scale process, whether it implements a logically reversible computation or anything else. These fluctuations disrupt their completion unless we employ entropically costly procedures to suppress fluctuations. The best we can achieve is a trade-off between improving the probability of successfully completing any single step and the increasing thermodynamic cost of achieving it. This "no-go" result for the foundations of the thermodynamics of computation has been derived in the Gibbs style in Norton (2013, Part 2; 2017) and in the Boltzmann style in Norton (manuscript, §5).

In its briefest form, the relation just is Boltzmann's celebrated relation "$S = k \log W$" between entropy S and probability W. If a thermodynamic process carries a system from state "1" to state "2," the driving entropy increase between the two states is $\Delta S = S_2 - S_1$. It relates to the probability of successful completion, W_2, and the probability that a fluctuation reverts the system to its initial state, W_1, by $\Delta S = k \log (W_2/W_1)$. Thus, a probability ratio favoring completion in molecular-scale processes can only be enhanced by a dissipative increase in entropy creation ΔS.

The outcome is that, independently of any entropy cost associated with erasure or the logic implemented, there is an inevitable entropy cost associated with the suppression of fluctuations. It depends only on the number of steps employed and the probability of completion sought.

This program-defeating neglect of fluctuations permeates the analyses of Bennett and Landauer. The program asserts that measurements are one of the logically reversible operations that can be carried out non-dissipatively, in direct contradiction with the foundational assumptions of Szilard and Brillouin. Bennett sought to illustrate non-dissipative measurements with a mechanical, rocking keel construction in (1987, p. 114) and one that uses variations in a magnetic field to enable the transfer of states between magnetic dipoles in (1982, pp. 929-31). If either process were attempted non-dissipatively, that is arbitrarily close to equilibrium, each would be disrupted by thermal fluctuations. As Earman and Norton (1999, pp. 13-14) argue, Bennett's keel would rock wildly and fail to measure anything.

Landauer has also offered schematic proposals for dissipationless measurements; and they fail in similar ways. The procedures he describes commonly involve steps that he incorrectly characterizes as dissipationless. That is, he describes them as involving "minimal energy requirements." Whatever he may have meant by that, each step is entropically costly, since each must suppress thermal fluctuations. For example, Landauer (1996, pp. 1915-16) describes a procedure involving coupled particles in time-dependent, bistable, potential wells.

The well configurations are manipulated by moving charges towards and away from the wells, which, he urges, "is not a source of energy dissipation." That is not so, since precisely this process is entropically costly in that its probabilistic completion, like that of all steps in the procedure, must suppress thermal fluctuations. The effect of this suppression is seen in the detailed computation of a related example in Norton (2013, §12).

Bennett, with Landauer's endorsement, has tried to establish that logically reversible processes can be implemented in thermodynamically reversible processes by positing idealized systems that achieve it. We can examine one to see how they fail. It is a most ingenious mechanical device that Bennett (1982, §3) calls a "Brownian computer." Its intricate mechanism implements a logically reversible computation largely through the motions resulting from "Brownian" thermal agitations. The mechanism is so designed that its state is restricted to a long, labyrinthine channel in a high-dimensional configuration space. Thermal agitation leads the system state to meander through this channel in a higher-dimensional analogy with Brownian motion. The computation completes when a system state falls into an energy well, the "energy trap."

Contrary to Bennett and Landauer's claim, it has been shown in Norton (2013a) that the device does not illustrate thermodynamically reversible computation. The free motion of the system state through its labyrinthine channel is, thermodynamically, equivalent to the uncontrolled, isothermal expansion of a single-particle gas in a high-dimensioned space. This is the standard example of a thermodynamically *irreversible* process. Such was the original treatment of Brownian motion in Einstein's (1905) celebrated paper. A thermodynamically reversible process is one in which all the thermodynamic driving forces are in near perfect balance. (For more see, Norton 2016.) In Brownian motion, the expansive, pressure force of the thermal motions is unbalanced so that the Brownian particle is unimpeded in a thermodynamically irreversible exploration of an ever-growing volume of its configuration space.

It is hard to see how Bennett could misidentify this iconic, irreversible process as reversible. His analysis tracks the *energy* dissipated in the process, such as that discharged as heat when the system state falls into the concluding energy trap. He should instead have tracked thermodynamic *entropy* directly, or its surrogate in these systems, *free energy*. An irreversible, entropy-increasing process may pass no heat to the environment, such as an unimpeded expansion of an ideal gas or its analog in Brownian motion. It was noted above that Landauer correspondingly sought to track thermodynamic dissipation in processes by surveying their energy requirements as opposed to their entropy creation. It is hard not to see efforts to assess thermodynamic dissipation by tracking energy as novice errors in thermodynamics.

Brownian computation does conform with the import of thermal fluctuation urged here. It is a simple affirmation that the fluctuation analysis is the better one. The analysis connects dissipation with the number of steps needed to implement a process. Accordingly, Brownian computers seek to minimize dissipation by reducing the entire computation to one long, slow step. The analysis urges that dissipative processes are needed to suppress the disruption of fluctuations. Accordingly, Brownian computers suppress fluctuations that would reverse the computation by introducing an entropically costly trapping of the system state in an energy trap.

Above, I have identified how the neglect of fluctuations has compromised the work of the founding authors of this literature. The neglect persists in recent work and remains troublesome. See, for example, Myrvold (2024) and Myrvold and Norton (2023).

4. Dynamic and Static Probabilities

Work in the thermodynamics of computation has long depended on conflating two senses of probability, the dynamic and the static. Probabilities "p" must be dynamic if the canonical entropy formula "$-k\ p \log p$" is to associate a thermodynamic entropy with a probability p. The probabilities of the thermodynamics of computation are static and, we shall see, are not generally connected with thermodynamic entropy.

The probabilities of statistical physics are dynamic. If we have some system that is in thermal equilibrium at temperature T, the probability that it has energy E is proportional to the Boltzmann factor, $\exp(-E/kT)$. These probabilities are realized by the system migrating dynamically in time over its accessible energy states, such that the frequencies approximate the probabilities. If the probability is uniform over some region of its configuration space, the system will migrate with roughly uniform occupation times throughout the region. Such a dipole will bounce back and forth uniformly between its up and down states. A one-molecule gas in a chamber with two accessible parts migrates repeatedly between them.

In the thermodynamics of computation, these same devices may be used as memory devices. A dipole in the "up" state may store a "1"; and one-molecule gas assuredly trapped in the left half of a chamber may encode "L." The associated probabilities are static. They represent our uncertainty over which state is stored in the memory. These probabilities are static. They are not associated with the dynamic migration of the memory device over its states. Such migration is precisely what a memory device cannot do, if is it to store information.

Early work in the thermodynamics of computation, such as Landauer's (1961) original paper, initiated the present practice by treating uncritically the static probabilities of such memory devices as if they had the same import as the dynamic probabilities of statistical mechanics. If a memory device can be in either disjoint state "0" or "1" with probabilities p_0 and p_1, it is assumed that the canonical entropy expression

$$S_{\text{info}} = -k(p_0 \log p_0 + p_1 \log p_1) \tag{1}$$

returns a thermodynamic entropy with the usual associations with heat. Under what is now called "Landauer's principle," erasure of the device is thermodynamically irreversible and, in generic circumstances, must pass heat TS_{info} to the environment.

This proposal fails as long as some general information-theoretic result is sought, in which the probabilities p can take any values. We shall see that it can succeed only when these probabilities are set to a unique set of values, fixed by the thermodynamics of the memory device. These are the probabilities arising in the thermalization of the memory device. The barriers that isolate the two memory device states 0 and 1 are dropped, so that the system can access both states dynamically. If V_0 and V_1 are the disjoint volumes of the *total* phase space associated with the two states, including the environment, then the unique, dynamic probabilities of the thermalized state are

$$p_{\text{dyn},0} = \frac{V_0}{V_0+V_1} \quad p_{\text{dyn},1} = \frac{V_1}{V_0+V_1} \tag{2}$$

That (1) can be recovered only with the specific probabilities (2) can be seen through a Boltzmann-style analysis of Norton (manuscript, §4). An erasure procedure, implemented by the same Hamiltonian for both states 0 and 1, evolves the memory device state to a reset state with total phase volume V_{reset}. Since Liouville's theorem requires the preservation of phase volume, the reset state must have a phase volume V_{reset} at least as great as the sum of the phase volumes. The resulting reset state can be treated as a single state if the evolved volumes, originating in V_0 and V_1, are combined by coarse graining. We have:

$$V_{\text{reset}} \geq V_0 + V_1$$

Since entropy S and phase volume V are related by $S = k \log V$, this process is dissipative, that is thermodynamically irreversible, and creates thermodynamic entropy. It effects the thermalization of the two states. The entropies of the states are:

$$S_{\text{reset}} \geq k \log (V_0 + V_1) \qquad S_0 = k \log V_0 \qquad S_1 = k \log V_1$$

If device states 0 and 1 arise with the static probabilities p_0 and p_1 of Landauer's analysis, it follows that the mean increase in entropy is:

$$S_{reset} - p_0 S_0 - p_1 S_1 \geq -k\left(p_0 \log \frac{V_0}{V_0+V_1} + p_1 \log \frac{V_1}{V_0+V_1}\right) \tag{3}$$

The result to note is that the expression on the right will only match the information entropy expression (1) if the probabilities p_0 and p_1 adopt the thermalization values (2).

It follows that Landauer's principle fails for an arbitrary set of static probabilities p_0 and p_1. The information entropy expression (1) does not, in general, recover the increase in thermodynamic entropy associated with erasure. It can only do so if the static probabilities are tuned to match the specific dynamic probabilities of a familiar process, thermalization. Once again, the elevated talk of a "principle" suggests novel results introduced by information and computation-theoretic thinking. Yet, in so far as the analysis goes beyond familiar results, nothing sustainable results.

This analysis explains an otherwise curious omission in the Landauer principle literature. In analyzing idealized memory devices with 0 and 1 states of equal entropy, no matter which static probabilities p_0 and p_1 are assigned to each state, the erasure procedures proposed always return an entropy cost of $k \log 2$. It corresponds to the maximum value of (1) arising when $p_0 = p_1 = ½$, which are the thermalization probabilities (2). The literature has offered no procedures that return a smaller entropy cost that would be associated by the information entropy formula (1) with other values of p_0 and p_1. This happens since all the erasure procedures available include a thermalization step that effectively overrides any probability distribution other than the dynamic probabilities (2) of $p_0 = p_1 = 1/2$. Since erasure is essentially thermalization, we now see that this eclipsing of other probability distributions is unavoidable.

A Boltzmann-style analysis precludes a general association between Landauer's information entropy and the increase in thermodynamic entropy during erasure. We should expect that a Gibbs-style analysis will return the same preclusion. The question must be handled with care, since Gibbs-style analyses can become convoluted and untenable assumptions may slip in unnoticed. Norton (manuscript, §§6-7) finds results that conform with the Boltzmann-style analysis. A careful Gibbs-style analysis does preclude such an association, as long as we require that a thermodynamic entropy has two properties: its changes in reversible processes correspond to Clausius' formula "heat/*T*"; and spontaneous changes correspond to increases in entropy.

This result strengthens the Boltzmann-style analysis, which only precluded an information-entropy term (1) corresponding to the *increase* in thermodynamic entropy that arises in the thermodynamically irreversible process of erasure. The Gibbs-style result applies to a compound state in which we are uncertain, with static probabilities p_0 and p_1, whether states 0 or 1 are present. It precludes identifying a thermodynamic entropy as a property of this compound state with the form of an information entropy (1). Were such an identification sustainable, erasure would be the passage of this quantity of thermodynamic entropy to the environment, so that the erasure process would conserve total thermodynamic entropy if the lower limit in (3) were realized. That would contradict the Landauer principle claim that logically irreversible erasure is thermodynamically irreversible.

Finally, it must be emphasized that the concerns of this section are overshadowed by more serious problems for the thermodynamics of computation. Its neglect of fluctuations has compromised much of its argumentation. The quantities of entropy creation it tries to attach specifically to computational processes are outweighed by those needed to suppress fluctuations.

5. Conclusion

The question of the relationship between thermodynamics, information and computation has created an enormous literature. Its volume is too great to be addressed here, in what must be a short paper. Critiques of the latest writing also risk dismissal as complaints about small lapses in minor additions to an already established literature. Instead, I have focused on the work of the founding figures of the field. My goal is to demonstrate that the field has been unsuccessful from the outset; and that the failings in the work of its founders persist. Present work continues in the thrall of the enticing idea that profound connections can be found between the concrete realm of steam engines and power stations and the abstract realm of information and computation. The connections, we have now seen, can only be sustained if they are trivialized by excising any content that goes beyond what are otherwise unremarkable results in existing statistical physics. More seriously, its literature is all too quick to dismiss a fatal problem: it neglects the controlling influence of fluctuations in molecular-scale processes as mere nuisances that perhaps someone else should resolve.

It has been so from the beginning of this literature nearly a century ago. It is so now and, I fear, that wishful thinking is powerful enough for it to remain so.

References

Bennett, Charles H. (1982) "The thermodynamics of computation—A review." *International Journal of Theoretical Physics*, **21**, pp. 905–940.

Bennett, Charles H. (1987) "Demons, Engines and the Second Law," *Scientific American*, **257** (No.5), pp. 108-117.

Bennett, Charles H. (2003) "Notes on Landauer's principle, reversible computation, and Maxwell's Demon," *Studies in History and Philosophy of Modern Physics*, **34** (2003) 501–510.

Brillouin, Leon (1951) "Maxwell's Demon Cannot Operate: Information and Entropy. I," *Journal of Applied Physics*, **22**, pp. 334-37.

Brillouin, Leon (1961) "Thermodynamics, Statistics, and Information," *American Journal of Physics*, **29**, pp. 318–328.

Carnap, Rudolf (1977) *Two Essays on Entropy*. Ed. Abner Shimony. Berkeley: University of California Press.

Earman, John and Norton, John D. (1998, 1999), "Exorcist XIV: The Wrath of Maxwell's Demon." Part I "From Maxwell to Szilard," Part II: "From Szilard to Landauer and Beyond," *Studies in the History and Philosophy of Modern Physics*, **29**(1998), pp.435-471; **30**(1999), pp.1-40.

Einstein, Albert (1905) "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen," *Annalen der Physik*. **17**, pp. 549–560.

ter Haar, Dirk (1966) *Elements of Thermostatics*. 2nd ed. New York: Holt, Rinehart and Winston.

Landauer, Rolf, (1961), "Irreversibility and heat generation in the computing process", *IBM Journal of Research and Development*, **5**, pp.183–191.

Landauer, Rolf (1991) "Information is Physical," *Physics Today*, **44**(5), pp. 23-29.

Landauer, Rolf (1996) "Minimal Energy Requirements in Communication," *Science*, New Series, **272**, No. 5270, pp. 1914-1918

Myrvold, Wayne C. (2024) "Shakin' All Over: Proving Landauer's Principle without Neglect of Fluctuations," *The British Journal for the Philosophy of Science*, **75**, pp. 587-616.

Myrvold, Wayne C. and Norton, John D. (2023) "On Norton's '…Shook…' and Myrvold's 'Shakin' …' " *Philosophy of Physics*, Vol. 1, Issue 1, Article 5.

Norton, John D. (2013), "All Shook Up: Fluctuations, Maxwell's Demon and the Thermodynamics of Computation," *Entropy*, **15** pp. 4432-4483.

Norton, John D. (2013a) "Brownian Computation is Thermodynamically Irreversible." Foundations of Physics. 43 (2013), pp 1384-1410.

Norton, John D. (2016), "The Impossible Process: Thermodynamic Reversibility," *Studies in History and Philosophy of Modern Physics*, **55**, pp. 43-61.

Norton, John D. (2017) "Thermodynamically Reversible Processes in Statistical Physics," *American Journal of Physics*, **85**, pp. 135-145.

Norton, John D. (2018) "Maxwell's Demon Does not Compute." in Michael E. Cuffaro and Samuel C. Fletcher, eds., *Physical Perspectives on Computation, Computational Perspectives on Physics*. Cambridge: Cambridge University Press. 2018. pp. 240-256.

Norton, John D. (manuscript) "The Simply Uninformed Thermodynamics of Erasure."

Rothstein, Jerome (1951) "Information, measurement, and quantum mechanics," *Science* **114**, pp. 171-175.

Smoluchowski, Marian (1912) "Experimentell nachweisbare, der üblichen Thermodynamik widersprechende Molekularphänomene," *Physikalische Zeitschrift*, **13**, pp. 1069–1080.

Szilard, Leo (1929) "On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings." Trans. Anatol Rapoport and Mechthilde Knoller. *Behavioral Science*, **9**(4), 1964, pp. 301-310.