In the 1830s, Charles Babbage worked on a mechanical computer he dubbed the Analytical Engine. Although some people around Babbage described his invention as though it had authentic mental powers, Babbage refrained from making such claims. He does not, however, seem to have discouraged those he worked with from mooting the idea publicly. This article investigates whether (1) the Analytical Engine was the focus of a covert research program into the mechanism of mentality; (2) Babbage opposed the idea that the Analytical Engine had mental powers but allowed his colleagues to speculate as they saw fit; or (3) Babbage believed such claims to be fanciful, but cleverly used the publicity they engendered to draw public and political attention to his project.

Charles Babbage (1791–1871) was the inventor of a number of mechanical devices that were intended to compute mathematical and logical functions. The two most famous are the Difference Engine, dating from the 1820s, which was to have calculated and printed basic mathematical tables of various kinds, and the Analytical Engine. The Analytical Engine, for which designs began in the mid-1830s, was a much more ambitious project because it was to have employed then-recent developments in abstract symbolic algebra, thereby extending its range into the logical realm rather than being limited to the strictly mathematical. Neither machine was ever completed, though Babbage built two small models of the Difference Engine.

For many decades now, it has not been uncommon to find Charles Babbage named in both popular and scholarly publications as one of the primary “forefathers” of computer science. This trend began with the public pronouncements of Howard Aiken, the designer of the Harvard Mark I “computer” in the 1940s, who repeatedly paid homage to Babbage and claimed to have been the one to finally fulfill “Babbage’s Dream,” as he often put it.

More recently, there has been a trend to cite Babbage as a father of compu-
tational cognitive science as well (e.g., Gardner, 1985; Garnham, 1987; Hauge-
land, 1985; Hofstadter, 1979; Pylyshyn, 1984). Though it has received little
serious consideration, this is a striking attribution, for it implies that Babbage not
only intended his machines to compute useful functions, but that he also thought
the operation of his machines in some way closely paralleled the operations of the
human mind. In this article I examine this claim. What I ultimately discover is that
Babbage himself was reluctant to associate the activities of his machines with the
operations of the mind, but he seems to have been happy to allow others, even
some who worked closely under his guidance, to come very close to attributing to
them actual mental powers. The question we are left with, then, is whether he did
this because (1) he was indeed engaged in covert psychological research, (2) he
was simply so liberal minded that he allowed even his closest associates to
publicly speak their minds even when their opinions of his work conflicted with
his own, or (3) he allowed his assistants to risk their own reputations by making
controversial claims in the hope that it would draw public attention to his project
which he could then use to his advantage.

The Origins of the Analytical Engine

Babbage had first made a name for himself in the late 1810s when he led a
mathematical revolution at Cambridge to switch from the clumsy Newtonian
notation for calculus to the more easily manipulated Leibnizian notation. For his
efforts, Babbage was made a fellow of the Royal Society at the age of only 24. His
“revolutionary fervor” was not spent, however. Frustrated with the stodgy and
aristocratic Royal Society, Babbage and a number of other “entrepreneur-scientists” (see Ashworth, 1994) formed the rival Astronomical Society in 1820. The
Astronomical Society began life by attacking the Royal Society-endorsed Board
of Longitude, which was then making up a set of new astronomical tables. While
busy discovering errors in the Board of Longitude’s tables (and just about
everyone else’s), Babbage hit upon the idea of developing a machine that would
calculate and print the required numbers automatically, so that there could be no
mistakes either in computation or copying. The machine he designed implemented
a then-common mathematical technique for calculating polynomials, from which
the values of a wide array of other functions could be approximated. The
technique was called the “method of differences,” and thus the machine was
dubbed the “Difference Engine.”2 He built a small working model in 1822, for

2 The method of differences allows complicated problems to be turned into simple addition
problems by relying on the fact that polynomials have certain constant relations to each other. For
instance, the square of the numbers 0, 1, 2, 3, 4, 5, . . . are 0, 1, 4, 9, 16, 25, . . . . By subtracting each
square from the square above it in the series (e.g., 1 – 0, 4 – 1, 9 – 4, . . .), one gets 1, 3, 5, 7, 9, . . . .
These are the “first differences” of the squares. If one subtracts each of these differences from
the one above it in the series (e.g., 3 – 1, 5 – 3, 7 – 5, . . .), the outcome is 2 in every case. These
are called “second differences.” Once one knows that the second differences between any square in
the series is 2, one can use the process in reverse to obtain a square simply by adding. For instance,
what is the square of 7? It is 2 more distant from the square of 6 than the square of 6 is from the
square of 5. The square of 6 is 2 more distant from the square of 5 than the square of 5 is from the
square of 4, and so on down to 0. Of course, this is a fairly cumbersome procedure for a human (who
can multiply), but for a machine that can only add (and that can do so extremely quickly and
accurately) it is a virtually ideal algorithm. Of course, figuring out the square of 7 is not going to
which he won the Astronomical Society’s “Gold Medal.” He then applied for assistance to build a full-scale version of the Engine to the Royal Society itself. Davies Gilbert, the Royal’s treasurer, lobbied the government on Babbage’s behalf. The Duke of Wellington, then Prime Minister, approved £1,000 for its construction in 1823 (Hyman, 1982, p. 52). Babbage found that the machine was more expensive to build than he had expected, so he asked for and received more money from the government, and then more, and still more. By 1833, a decade later, he had received over £17,000, and still no working machine had been completed (but for a new scale model that he kept in his house as a showpiece). The government finally suspended the project in the mid-1830s, officially withdrawing their support in 1842. Then-Prime Minister Robert Peel is reported to have quipped that, if ever completed, the Difference Engine should first be set to calculating how much money went into its construction.

In the late 1820s, Babbage toured Europe, devoting much of his travel to studying Continental means of industry and manufacture. One of the things he discovered was the system set up by Marie de Prony (1755–1839) in Revolutionary France for the calculation and publication of new logarithmic and trigonometric tables. Faced with this monumental task, De Prony had innovatively applied Adam Smith’s (1723–1790) analysis of the division of labor (Smith, 1776, chap. 1), developed for traditional manual work, to what was a mental task. He established three sections of mental laborers: The first section was to consist of the best mathematicians in the land, who would seek out or develop the easiest methods of calculating the functions desired; the second section, consisting of six or eight junior mathematicians, would take the formulae developed by the first section and plug in the actual numbers required; the third section, made up of 60 or 80 people who could only add and subtract, would do the actual computations and return the results to the second section for checking. Babbage, in chapter 20 of his 1832 book *On the Economy of Machinery and Manufactures*, hints that his Difference Engine could replace the whole third section of laborers in De Prony’s system. It is notable, however, that throughout the chapter the emphasis is on efficiency and economy; the nature of mental processes themselves is never at issue. As we shall see, Babbage seems to have assiduously and subtly dodged the question of the nature of the mind throughout his career.

It was in 1833 that he first met the future Ada Lovelace, then known as Ada Byron, the only legitimate child of the poet Lord Byron. Then aged 17, Ada was enjoying her first London season. Among other things, Babbage was well known for his Saturday soirées, and Lady Byron and her daughter were invited along with much of London Society. Babbage’s scale model Difference Engine was displayed along with other mechanical devices of a more obviously entertaining nature. Lady Byron recorded in her diary that in mid-June 1833 they went to see what she called “the thinking machine” at Babbage’s house (cited in Stein, 1985, p. 42). This is the first of many references to Babbage’s machines actually having mental powers. What Lady Byron’s remark shows is that even at this early date,
before the plan for the more sophisticated Analytical Engine had even been developed, the idea was already “in the air.”

During the middle and late 1830s, Babbage began developing a design for the Analytical Engine. Unlike the Difference Engine, which could only calculate functions reducible to the method of differences, the Analytical Engine would be able to calculate any function whatsoever. Indeed, he often claimed that ultimately it would be able to carry out symbolic algebra, then considered to be one of the highest forms of rational thought known. To accomplish this extraordinary task, Babbage adapted the technology developed by Joseph-Marie Jacquard (1752–1834) in his famous automated loom of 1801: He would lash cards together end to end and punch holes in them that could be “read” by a number of movable pins in the machine. The holes would “encode,” as we would now say, information about which operations to employ over which symbols. Given a sophisticated enough control system for the cards, one would be able to repeat the same set of cards an indeterminate number of times (now called “looping,” but he called it “backing” the cards), and one would be able to decide which cards to execute on the basis of intermediate results obtained during the computational process (now known as “conditional branching”). Both processes are central to modern computing theory.

Babbage attempted to obtain support for his new project but found it impossible to do so. Few understood the potential of the new invention. Few of those who did thought it would be practicable to build a machine so complex. Fewer still believed that Babbage was the man to finish the project, given the fiasco with the still-unfinished Difference Engine. Babbage began on his own nevertheless. In 1840 he was invited by the well-known Italian scientist Giovanni Plana to give a series of lectures on the Analytical Engine in Turin. An article based on the lectures was to be written by a then-unknown Italian military engineer, Luigi Frederico Menabrea (1809–1896). (Menabrea would later go on to become a general in Garibaldi’s army, and eventually Foreign Minister and Premier of the newly unified Italian state.) Babbage’s hand was never far from the article written under Menabrea’s name. He spent a good deal of 1841 corresponding with Menabrea about the article, even prompting the young Italian to entirely rewrite his article at one point (Hyman, 1982, p. 190). It appeared in French in the Swiss journal *Bibliotheque Universelle de Geneve* in 1842. In February of 1843, Charles Wheatstone—coinventor of the telegraph and a friend of Babbage—suggested to Ada, now Lovelace, that she produce an English translation of Menabrea’s article. According to Babbage, when Lovelace presented him with the completed translation, he suggested that she write some notes explicating parts that had been since superseded by new developments. As with the original Menabrea article, his involvement in the writing of the notes was significant. The translation and notes appeared in the September 1843 issue of Richard Taylor’s *Scientific Memoirs*—an English journal that specialized in communicating Continental European scientific activities to the British scientific community (Menabrea, 1961).

Ada Lovelace’s translation of and notes to Menabrea’s article did not have the impact that Babbage must have hoped they would. Although Lovelace was a bright person, to be sure, she was not a serious scientist or mathematician. She had never before published a scientific article, nor would she publish another in the remaining 9 years of her life. There can be little doubt that Babbage’s main goal
Was the Analytical Engine Thought to Be Cognitive?

As mentioned above, in recent years Babbage has been frequently put in the role of “trailblazer” in books and articles not only about computer science, but in those about cognitive science as well. The question is, then, whether this status is really justified. Were either Babbage, Menabrea, or Lovelace of the opinion that the Analytical Engine really would have been able to think—like humans think—if it had been completed? In this half of the article I try to show that Babbage seems to have been wary of answering this question straightforwardly throughout his life. Menabrea denied the possibility outright. Lovelace often wrote as if she believed it to be true, but at other times made what appear to be forthright denials of the possibility. Given that Babbage had so much influence over her writing, it is interesting to speculate about why he allowed her to occasionally stray over this line if he was unwilling to do so himself. One reason that he restrained himself may have been, or so I will argue, that for him to have done so would have made him appear to endorse mechanistic materialism, which would likely have been tantamount professional suicide in early Victorian England (see, e.g., Winter, 1997). But if this was his fear, then his reasons for his not dissuading Lovelace from doing so remain obscure. First I will examine the Menabrea/Lovelace paper for indications of the kind of cognitive theory, if any, that is contained therein. Then I will survey Babbage’s comments on the topic and offer an interpretation of them.

Menabrea opens his paper by claiming that the tasks of mathematics may be divided into two parts, “one of which may be called the mechanical, . . . while the other demanding the intervention of reason, belongs more specially to the domain of the understanding” (p. 225). Machinery, he went on almost analytically, may be employed to execute the mechanical portion of mathematics. After a brief mention of Blaise Pascal’s mechanical adding machine, Menabrea then reviewed the portion of Babbage’s Economy of Machinery in which De Prony’s scheme for producing mathematical tables was discussed, and stated outright that the third section of workers could be replaced by the Difference Engine. Here Lovelace inserts her first note. She first takes pains to deny that the Analytical Engine has any relation whatever to the Difference Engine—a claim that Babbage needed to make stick if he were to have people take seriously his pleas for support for the new project. She then goes on to describe the difference between them as follows:

In studying the action of the Analytical Engine, we find that the peculiar and independent nature of the considerations which in all mathematical analysis belong to operations, as distinguished from the objects operated upon and from the results of the operations performed upon those objects, is very strikingly defined and separated. . . .

It may be desirable to explain that by the word operation, we mean any process which alters the mutual relation of two or more things, be this relation of what kind

3 All page references for this article are to the reprint found in Morrison and Morrison (1961).
it may. . . . In abstract mathematics, of course operations alter those particular relations which are involved in the consideration of number and space. . . . But the science of operations . . . is a science of itself and has its own abstract truth and value; just as logic has its own particular truth and value, independently of the subjects to which we apply its reasonings and processes. (pp. 247–248)

Here we have a fairly standard formulation of the new formal approach to algebra then being pioneered by Lovelace’s calculus teacher, Augustus De Morgan. Lovelace likely learned this orientation toward algebra from De Morgan himself and from the textbook he assigned to her, George Peacock’s *Treatise on Algebra* (1803), in which the revolutionary distinction was first made between “arithmetical algebra”—the traditional manipulation of numerical expressions via symbols—and “symbolic algebra”—a new general discipline of abstract symbol manipulation, of which the arithmetical form was but a single application.

Lovelace then moves on to exemplify her claims about abstract algebra:

Supposing, for instance, that the fundamental relations of pitched sounds in the science of harmony and of musical composition were susceptible of such expression and adaptations, the engine might compose elaborate and scientific pieces of music of any degree of complexity or extent.

The Analytical Engine is an *embodiment of the science of operations*. (p. 249)

Today this passage is often mistaken for the claim that the Engine would be “intelligent” enough to compose music, but that is not really its thrust at all. The aim is to drive a wedge between abstract algebra—the “science of operations”—and mathematics; to show that the operations can operate over symbols representing objects other than numbers, such as musical notes.

After describing the differences between the two Engines even more fully, she then makes a claim that might seem to virtually commit her, and perhaps Babbage, to mechanistic materialism with respect to the mind (very close to the time that Helmholtz, Brücke, Bois-Reymond, and Virchow were declaring themselves to be mechanists in Berlin, Germany):

In enabling mechanism to combine together general symbols in successions of unlimited variety and extent, a uniting link is established between the operations of matter and the abstract mental process of the most abstract branch of mathematical science. . . . We are not aware of its being on record that anything partaking in the nature of what is so well designated the Analytical Engine has been hitherto proposed, or even thought of, as a practical possibility, any more than the idea of a thinking or of a reasoning machine.

Here she claims outright that the Analytical Engine serves as a “uniting link” between mind and matter, and then says that no one else had yet proposed a machine like the Analytical Engine any more than they had proposed a thinking machine. Although she stops somewhat short of the explicit claim that the Analytical Engine is itself such a thinking machine, the clear intent appears to be to raise the possibility for serious discussion.

Returning to Menabrea’s original text, having just stated that the Difference Engine could replace De Prony’s third section of workers (those who actually carry out the calculations) he goes on then to virtually claim that the new
Analytical Engine would be able to replace the second section as well (those who plug the numbers into the formulas produced by the expert mathematicians in the first section). He is careful to point out, however, that the machine “must exclude all methods of trial and guess-work, and can only admit the direct process of calculation.” This is necessarily the case, he goes on to say, because “the machine is not a thinking being [italics added], but simply an automaton which acts according to the laws imposed upon it” (p. 230). This would seem to close the matter definitively, but Lovelace adds a footnote here, saying that “this must not be understood in too unqualified a manner. The engine is capable, under certain circumstances, of feeling [italics added] about to discover which of two or more possible contingencies has occurred” p. 230). What exactly she meant by this is not entirely clear, but she appears to be referring to the Engine’s conditional branching ability. Why she took this to manifest itself as a “feeling” is anyone’s guess, though Babbage’s occasional picturesque references to this power as “foresight” may be the source.

So already we can see a tension between two different ways of viewing the machine. On the one hand, according to Lovelace, it “unites” matter and mentality. She cannot quite bring herself to calling it a “thinking machine,” but she’s willing to say that it “feels” in some sense. On the other hand, according to Menabrea, it is definitely not a “thinking being,” just an automaton. Under other circumstances one might regard this simply as a difference of opinion, but since both these writers had one and the same mentor—Charles Babbage—a more interesting dynamic may be at play.

Menabrea says nothing more about the possible relation between the operation of the Analytical Engine and that of the human mind in the rest of his article, but Lovelace includes one more major speculation in her notes—the passage that Alan Turing (1950), over a century later, would dub the “Lovelace Objection” to artificial intelligence:

It is desirable to guard against the possibility of exaggerated ideas that might arise as to the powers of the Analytical Engine. In considering any new subject, there is frequently a tendency, first, to overrate what we find to be already interesting or remarkable; and, second, by a sort of natural reaction, to undervalue the true state of the case, when we do discover that our notions have surpassed those that were really tenable.

The Analytical Engine has no pretensions whatever to originate anything. It can do whatever we know how to order it to perform. It can follow analysis; but it has no power of anticipating any analytical relations or truths. (p. 284)

If there were ever a question of Lovelace’s believing that the machine would actually be able to think, this passage would seem to scotch it—certainly Turing thought so—but in combination with her earlier statements about it unifying mind and matter, and its being able to “feel,” it is hard to know exactly what opinion to attribute to her, or indeed to her mentor, Babbage.

Babbage’s comments on the topic were always very careful. He almost invariably referred to the machine’s activities as being able to “replace” or “substitute for” mental activities, or occasionally as being “analogous” to them, but only very rarely as being able to carry them out itself. For instance, in his
autobiography (Babbage, 1864/1994), in the midst of a discussion of his difficulties in mechanizing the process of “carrying” from one column of numbers to another in addition problems, he says, “the mechanical means I employed to make these carriages bears some slight analogy to the operation of the faculty of memory” (p. 46). With this deflationary statement in place, he says later on the same page that “it occurred to me that it might be possible to teach mechanism to accomplish another mental process, namely—to foresee” (p. 46). He then goes on to explain what we would call conditional branching.

What are we to think of Babbage’s “true” beliefs on the matter? Was the first comment about memory a mere “foot in the door” that would allow him to slide to the stronger claim about foresight, or is the claim about foresight to be regarded as a mere “shorthand” for his real position, given just above it, that the activities of the machine are only “slightly analogous” to mental activities? Babbage seems to have clarified his position in a paper entitled “On the Mathematical Powers of the Calculating Engine” that had been authored back in 1837 (but that went unpublished until 1973):

In substituting mechanism for the performance of operations hitherto executed by intellectual labor it is continually necessary to speak of contrivances by which certain alterations in parts of the machine enable it to execute or refrain from executing particular functions. The analogy [italics added] between these acts and the operations of mind almost forced upon me the figurative [italics added] employment of the same terms. They were found at once convenient and expressive and I prefer continuing their use rather than substituting lengthened circumlocutions.

For instance, the expression “the engine knows, and so forth” means that one out of many possible results of its calculations has happened and that a certain change in its arrangement has taken place by which it is compelled to carry on the next computation in a certain appointed way. (Babbage, 1989, Vol. 3, p. 31)

Thus, it seems to have simply been a matter of verbal shorthand for Babbage, not a literal claim of mentality.

One other interesting passage occurs in the posthumous biography of Babbage written, at his direction, by his long-time friend Harry Wilmot Buxton (not published until 1988). It begins as though a profound claim concerning the intelligence of the Engine is about to be defended:

It is manifest that the language of algebra is more simple and precise than the symbols of language expressed by sound, and it would seem therefore within the range of our intelligence to be able to reduce our thoughts in most cases into the form of mathematical language, and thus adapt the subject of our enquiry to the operations of the Analytical Engine. (Buxton, 1988, p. 155)

Buxton then goes on, however, to make a rather vague claim about Babbage’s aspirations in this regard—“Mr. Babbage entertained no doubt of the possibility of extending the powers of the Analytical Engine, far beyond the domain of abstract analysis” (p. 155)—but then he immediately shifts to a long quotation from Hobbes in which it was claimed that reasoning is nothing more than calculation. Buxton next suggests that the Analytical Engine could have established Hobbes’ claim if only it had been equipped with letters of the alphabet and
the plus and minus signs, in addition to numerals. Having gotten this far, however, he closes with the admission that “as Mr. Babbage himself has not recorded his views upon the subject, it might be deemed presumptuous to indulge in speculations or enter into details which he did not deem necessary” (Buxton, 1988, p. 156). So it appears that even Babbage’s closest friends were not of the opinion that he believed his machines to be literally intelligent.

**Conclusion**

In this article I have attempted to establish that, contrary to what is often found in textbooks, Babbage did not believe the Analytical Engine to have been a contribution to what we would now call cognitive science. Why might Babbage have refrained from taking such a step? Such speculation was rampant about him: Lady Byron had called even the lowly Difference Engine a “thinking machine”; Buxton certainly entertained the idea that human reasoning might be nothing more than what the Analytical Engine was intended to do; even Ada Lovelace intermittently raised the possibility. What held Babbage back?

Two things come immediately to mind. The first is that in Regency and early Victorian England, declaring oneself to be a mechanist materialist with respect to the mind would have been professionally foolish in the extreme. Alison Winter (1997) has written a fascinating chapter on the philosophical traps of this sort that the famed physiologist William Benjamin Carpenter—who, incidentally, briefly served as tutor to Ada Lovelace’s children—had to avoid in launching his career during the 1830s and 1840s; dangers that, for instance, the notorious mesmerist John Eliotson had failed to negotiate effectively, leading to his resignation from University College in 1838.

Tempting as it might be, however, one should not conclude that Babbage somehow “secretly” believed the Analytical Engine to be truly intelligent, but would not say so publicly for fear of censure. Babbage, though liberal, appears to have been a sincerely religious man. When William Whewell published in 1833 the first in a series of works commissioned to defend the thesis that God’s “design” can be discovered in the organization of the natural world, Babbage felt compelled to write a book-length response, his so-called *Ninth Bridgewater Treatise* (1837). Significantly, Babbage’s objection to Whewell was not that he defended “natural theology,” as it was then called, but rather that Whewell had argued that nothing of the divine could be found in the “deductive” sciences, but only in “inductive” forms of knowledge. Babbage was incensed, not only because this excluded the truths of mathematics from the realm of the “divine”—a position Babbage was keen to defend—but also because the intent of Whewell’s remarks was to devalue the work of recent Continental scientists (e.g., Euler, Laplace, Lagrange). Babbage thought these to be the very examples to which the English should look to improve their science (Richards, 1992, pp. 60–61). Although Babbage argued that miracles might be the result of a “divine mechanism” the underlying principles of which we humans can only partially comprehend, his suggestion was not that God Himself is such a mechanism; it was, by contrast, that only a truly divine intelligence could craft a mechanism so intricate as that found in nature.

To conclude, then, Babbage does not seem to have regarded the Analytical
Engine as a kind of intelligence, or even as revealing anything of particular significance about the nature of the human mind. Attributions of such opinions to him by modern cognitive scientists and their intellectual kin are primarily anachronistic. His primary interests in his Engines were industrial and economic. He saw them as bringing the very same principles of division of labor to the realm of mental work that he supported in the realm of manufacture. Indeed, one might argue that the powers of his Engines served to distinguish precisely between those aspects of the human mind that were thought to be merely mechanical and those that were regarded as being truly original, creative, rational, and ultimately divine in character.

References


Correction to Kosits (2004)


On p. 358, the first paragraph should read “Given this theological background, it is now possible to consider the New Psychology’s Fallacy argument as deeply ironic and even tragic, particularly for William James, the argument’s most influential articulator.”

On p. 342, footnote 5, 6th line of the quotation, the word to should not be crossed out.