There and Back Again?
Reproducibility and the Hunt for a Human Compass Sense
by
Michael A. Greshko
B.A. Biological Sciences, Vanderbilt University (2014)
SUBMITTED TO THE PROGRAM IN COMPARATIVE MEDIA STUDIES/WRITING IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN SCIENCE WRITING AT THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY
©2015 Michael A. Greshko. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature redacted
Signature of Author: ____________________________ Michael A. Greshko
May 21, 2015

Signature redacted
Certified By: ____________________________ Alan Lightman
Professor of the Practice
Thesis Supervisor

Signature redacted
Accepted By: ____________________________ Thomas Levenson
Professor and Director, Graduate Program in Science Writing
There and Back Again?
Reproducibility and the Hunt for a Human Compass Sense

by

Michael A. Greshko

Submitted to the Department of Comparative Media Studies/Writing on June 3, 2015, in Partial Fulfillment of the Requirements for the Degree of Master of Science in Science Writing

ABSTRACT

Living creatures must navigate their environments in search of food, reproductive opportunities, and better habitats, and they use many stimuli in order to do so. After centuries of skepticism, biologists in the 1960s convincingly demonstrated that the Earth’s weak, omnipresent magnetic field was also detectable by animals trying to orient themselves in space, a sense dubbed magnetoreception. Long enchanted with animal migration, University of Manchester biologist Robin Baker asked a fateful question: Why not humans? From 1976 to the late 1980s, Baker amassed evidence that he claimed as proof that humans had a magnetic homing sense. When Baker’s experimental subjects were blindfolded and displaced in a variety of settings, they could orient better than chance toward their original location or along assigned compass directions. Subjects wearing magnets on their heads, however, could not.

Problematically for Baker, his peers were largely unable to replicate his results, leading to a passionate academic debate that lasted throughout the 1980s. His critics lambasted him over issues of experimental design, unconscious bias, and statistical false positives, while Baker accused his critics of misrepresenting their own data. Having exhausted his interest in the field—and undoubtedly weary of the challenges to his work—Baker stopped studying magnetoreception in the late 1980s, though he stands by his claims to this day. No researcher since has taken up the question of human magnetoreception with similar commitment, and Baker’s results have remained controversial and largely unaccepted by the larger scientific community.

Baker’s case illustrates the necessity of reproducibility in science and underscores science’s messy realities, a point similarly shown by controversial incidences of “pathological science,” including Blondlot’s discovery of N-rays, Weber’s detection of gravitational waves, and Fleischmann and Pons’ announcement of cold fusion. Baker’s pursuit of the human magnetic sense also provides insight into the importance—and potentially self-deceiving dangers—of passion as a motivating force for scientists.
Acknowledgements

Many thanks to:

My thesis adviser Alan Lightman, for his wry sense of humor, his charity of spirit(s), and knack for pointing me in the right direction;

My second reader Marcia Bartusiak, for her insightful feedback;

The indomitable faculty of the Graduate Program of Science Writing, for giving me a marvelous toolkit I hope to use for a long time to come;

Shannon Larkin, for her incredible support through the good, the bad, and the ugly (e.g. the gut-wrenching typo on the cover of this document);

My fellow Border Collies—Rachel Becker, Chris Couch, Cara Giaimo, Anna Nowogrodzki, Sarah Schwartz, and Joshua Sokol—for the opportunity to learn with and from them;

My family, friends, and loved ones, for putting up with my shenanigans;

And my sources, particularly Robin Baker, for their generosity and time.
Robin Baker has lost his voice.

The former University of Manchester lecturer has a nagging throat condition, he tells me, so we are forced to fall into a written rhythm. I send him batches of six emailed questions, and in a few days’ time, he answers point by point, his blue paragraphs doing a respectable job of approximating spoken cadence. But beyond his words, it’s hard to get a read on him. I can’t scrutinize his face or hear his voice. For that, I have to resort to old TV spots he’s done through the years, like an October 1982 episode of Naturewatch:

It’s a sunny day in southwestern England, and Robin Baker is in his element. Amid the murmur of the nearby birds and throngs of undergraduates standing before him, the University of Manchester biologist seems to savor the sun, the grass, the tree branches moving gently in the wind. He’s leaning on a rock wall, and between his soothing voice, thick, dark brown beard, and blue-shirt-and-jean combo, he’s suaveness incarnate—the Barry Gibb of Science. He begins to speak, his shoulder-length brown hair blowing slightly in the breeze, and all eyes turn to him.

“This morning, we’re going to do what we call a walkabout experiment,” he says, gesturing all around at Woodchester Park, where they’ve gathered for the day. “We take you on a walk through the woods here.” He gestures north, south, east, and west, as a reference for questions later on; throughout, his audience is giggling. He’s utterly charming.

It’s only then that he gestures to the electrified helmets he wants them to wear.

Cut to voiceover. “There’s no doubt in my mind that humans have a magnetic sense. We’ve got enough data, I think, to make it clear that a magnetic sense exists and is being used.”

He cheerfully leads his subjects out on a walk.

“The challenge and the fun is finding out just how it’s being used—and why it sometimes doesn’t work.”

I’d love to see him in an unguarded moment like that walk in the park. But that can’t be. In fact, I wonder if the illness keeping us confined to email is simply a means of controlling the situation. “I always like to have a record of what I have said,” he reminds me well into our correspondence.

It helps that his prose is lively. Since leaving England for southern Spain in 1996 with his partner and three children in tow, Baker has made a second career of it, publishing a variety of saucy pop-sci books and novels outlining his ideas on human sexuality: Sperm Wars. Sex in the Future. Baby Wars. Primal. He’s got a knack for this sort of thing, gaining notoriety in the 1990s for his revolutionary—and largely panned—explanations for the shape of the human penis, the function of the female orgasm, and how sperm from different men “fight” one another in the vaginas of promiscuous women.

Before studying and writing about sexuality, though, Baker had a first academic love: animal navigation. From homing pigeons’ incredible ability to fly home to sea turtles’ near
miraculous return to their beach birthplaces, people have been fascinated since time immemorial by how animals—including humans—figure out how to move through the world. It’s a complicated problem that animals address using a litany of tools. Sight, smell, hearing, the location of the sun and stars in the sky, you name it—if animals can sense it, chances are that they use it to help them find places to eat, live, and breed.

By the time Baker first began studying animal migration and navigation in depth, scientists had just discovered another tool to animals’ navigation toolbox: “magnetoreception,” the ability to orient using Earth’s magnetic field as a guide. Long dismissed as fringe science, the hunt for this compass-like “sixth sense” finally paid off in the 1960s and 1970s, with clear evidence beginning to pile up in everything from the birds to the bees. Struck with inspiration, Baker asked a fateful question in the late 1970s: what about humans?

In a flurry of tests culminating the 1980 publication of a study in the super-journal *Science*, Baker claimed to show that disoriented people could point homeward in the absence of visual cues. What’s more, his results showed that disrupting the magnetic field surrounding the head could disrupt this ability to orient. Like other animals, humans could detect magnetic fields, it seemed. “And that’s when I realized that for whatever reason, there was a lot of base hostility among the academic community,” wrote Baker in his first email to me.

The community’s problem with Baker was that Baker seldom had company in getting the results he was claiming. Throughout the 1980s, researchers in the United States and Europe worked to replicate his results but hardly ever could, leading many to doubt Baker’s conclusion that humans, in fact, have a magnetic sense. The debate was bruising. After a decade of academic infighting that spanned many papers and three books, Baker eventually turned away from the field that had turned away from him, moving on to study human sexuality after his last major work on the topic in 1989.

Baker’s search of the human magnetic sense and the resultant backlash illustrate one of the most important yet unsung components of science: reproducible experimental results. While the process of replicating a researcher’s work can be messy, vindictive, and thoroughly political—a sort of back-and-forth that can lead to tension, name-calling, and outright hostility, as in Baker’s case—the scientific method forces us to understand the natural world through the development of repetition and growing consensus among entire communities of researchers. Whether or not these replication efforts are successful determine if even the most brilliant ideas live or die.

*If It’s Worth Doing...*

In science, making a claim like Baker’s is an inherently bold thing to do. The general public may be lulled into a sanitized view of science, a pile of textbooks dug up by some distant Scientists. But every time researchers publish a result, they’re describing a facet of Nature that no one else has ever seen, opening themselves up to a dynamic community clamoring to test out—and sometimes shout down—new evidence and ideas. Science, in a word, is messy.
Science historian Thomas Kuhn famously described this process as a series of community-wide revolutions fueled by high-octane, reproducible results. According to Kuhn, entrenched scientific worldviews—which he termed "paradigms"—guide new research, as scientists work to figure out the finer details of current theory. But sometimes, experimental results accumulate that don’t quite fit with the old models. Initially, there’s considerable skepticism: perhaps these outlier experiments are merely the products of mistaken researchers.

However, if others can replicate these inexplicable results, then concern over the anomalous results evaporates. Excitedly, corners of the scientific community begin working on new theories, in an effort to incorporate these strange new results. And if these new explanations stick, then the "paradigm shifts": The scientific community jumps ship and rapidly moves over to the new explanation. And so the cycle continues.8

The real key is that initial phase, where individual scientists publish papers that challenge old paradigms. Group-wide skepticism is a healthy response to apparently bizarre results, precisely because it spurs curiosity—and the desire to check others’ work. If that never happened—if extraordinary claims were simply accepted at face value—then scientific theory would be rendered unstable, constantly shifting with every new paper being published. Hucksters and frauds could go unnoticed, slipping by with made-up results. The pile of textbooks would be rendered a house of cards.

But for all their merits, the odds are understandably stacked against replication studies. Frankly, it’s dull, thankless work. Scientists often would rather seek answers to their own questions than spend time and money double-checking their peers, particularly in a day and age where there are so many original questions to ask. Scientific journals, keen to keep up their academic sex appeal, like publishing snazzy, new research—not rehashes of previous work. And unless a replication attempt contradicts a study that got major media airtime, news outlets rarely have an interest in covering them.

Science, however, needs replication studies now more than ever. In 2005, epidemiologist John Ioannidis published a watershed essay in the journal PLoS Medicine, persuasively showing through statistical argument “why most published research findings are false,” a result of false positives.9 The article has been viewed nearly 1.3 million times since, read and reread by a scientific community increasingly made skittish by results that haven’t been holding up. In 2011 and 2012, for instance, pharmaceutical giants Bayer and Amgen announced that they could only replicate a handful of preclinical studies on cardiovascular, cancer, and women’s health issues—a worrying finding echoed by University of Texas researchers in 2013.10,11,12 Why so many issues with reproducibility? Troublingly, it’s hard to say. Complicated procedures? Sloppiness?

In at least one high-profile case in the last year, the answer is unnerving: outright fraud. In January 2014, a Japanese researcher named Haruko Obokata authored two papers that were published in the journal Nature. Combined, Obokata’s work seemingly pointed to a shockingly simple way to make stem cells, highly malleable proto-cells that can transform into different tissues. This was a huge deal: Stem cells are ideal for treating all sorts of diseases, but researchers either have to harvest cells from embryos—opening up ethical quagmires—or have to coach adult cells to revert to their original, versatile states.
But Obokata claimed that simply dunking adult cells in acid baths could stress them enough to become stem cells. If true, the discovery stood to change the nature of stem cell research, possibly accelerating the development of new medical treatments. Almost immediately, however, researchers noticed problems with the study: Nobody could reproduce Obokata's claimed effect. Soon, Japanese authorities launched an investigation, and by April, it found Obokata guilty of scientific misconduct. By year's end, the Nature papers were withdrawn, and Obokata had resigned from her position.

Irreproducible results like Obokata's fraud is one thing; Baker's eventual path toward the margins was altogether different, as we will see. In particular, Baker was spurred on by immense intellect and love of advocating for his opinions, often to the point of obstinacy—a habit he picked up in his career before pursuing the magnetic sense.

**The Happy Warrior**

Born on March 13, 1944 in Wiltshire, England—about twenty-five miles north of Stonehenge—Robin Baker was slightly unlucky from the very beginning. Born a mere four minutes after midnight, Baker's mother "was so dismayed at giving birth on the 13th by so narrow a margin that she asked the midwife to record the date of birth as the 12th," as Baker recounted on his impressively detailed online autobiography. The data, however, couldn't be questioned. The midwife flatly refused.

He entered the world the youngest of four children, his father's second child and his mother's third. Both parents worked: His father, a WWII veteran, maintained the village's manor, while his mother worked as a nurse in the local mental hospital. His childhood was largely happy, spending much of his time with his older sister and half-sister running about Manningford Bruce, the small village in Wiltshire that the family called home. Between playing with the dog and marveling at butterflies—his own "'Cider with Rosie' existence," writes Baker—he loved spending time outside, running amok in his idyllic, quaint hometown. But it wasn't just all play. Until 1955, he attended the local primary school in Manningford Bruce, a "tiny 2-room" affair with fewer than 30 students.

Once Baker turned 11, he was off to the nearby Marlborough Grammar School, where Lord of the Flies author William Golding had attended school three decades before. Graduating in 1962—and walking away with two high school tennis championships—Baker made the jump to the University of Bristol, where he finally could indulge his childhood love of butterflies. Finishing off his bachelor's degree in 1966, he stayed on to complete a doctorate in 1969, spending his time studying the intricacies of how butterflies flit and flutter from place to place.

Baker took to his PhD adviser, Howard Hinton, like a hungry monarch to milkweed. A tireless and innovative researcher who managed an active lab, edited multiple academic journals, and chaired the University of Bristol's biology department, Hinton was a force to be reckoned with. Baker and Hinton seem to have gotten along well during their time together. A Fellow of the Royal Society, Great Britain's most prestigious scientific organization, Hinton shepherded
Baker's first scientific publication into its records, a treatise exploring the evolution of butterfly migration.23,24

But for all that Hinton groomed Baker and honed his scientific skills, Baker might have picked up a couple of bad habits from him. Hinton's passion—manifesting itself in the form of off-color jokes or unabashed denouncements of his peers' scientific views—often clouded "whether he was stating a fact that he had established or a firmly held opinion," as his official Royal Society biography tactfully put it. A proud stirrer of the pot, Hinton carried "a reputation as a scientific polemicist" and "rather enjoyed that view of himself."25

Baker first exhibited a similar love of sparring, aided and abetted by a brilliant scientific mind, while a young lecturer—first at the University of Newcastle-upon-Tyne, and then the University of Manchester, in 1976. It was a time for Baker to flex his creative, confrontational muscles. In his view, As biologist Tim Birkhead, a student of Baker's at the time, later wrote: "Robin loved controversy...[He] had novel ideas and sold them hard—but at a cost."26 It turns out that Baker "had been scorned by other insect biologists" for his expansive views on migration, which were far more generous than those of his peers.

At the time, migration—a phenomenon primarily studied and defined by ornithologists—was seen in the shadows of what birds did: the regular mass movement of individuals from one place to another and back again, usually timed with the seasons and tied to breeding. As an entomologist studying what he saw as butterflies' migration, however—a looser mass movement from one place to another—Baker thought that other animal movements also qualified. He was heretically general. What if all these different animals' motions were facets of the same grand design?

In a flash, Baker saw an opportunity for a vast synthesis, one that encapsulated all the ways in which animals moved from one place to another. Fresh off a book deal with British educational publisher Hodder and Stoughton that allowed him to do "anything [he] wanted," he set to work "trying to tie together the movements, orientation and navigation of all animals, from plankton to humans," diligently writing over the next two years to compile a massive volume entitled The Evolutionary Ecology of Animal Migration.27 The ornithologists' narrow-minded views of migration were simply misguided, Baker thought. He had "become convinced that there were underlying mechanisms across all groups," he said in an email to me, and was going to prove it—by any word count necessary. Over more than a thousand pages, Baker argued that all animal movements, from "the small-scale...daily movements of aphids is the same phenomenon as the seasonal movements of trout."28 No matter the movement, no matter the distance, it was all "migration": utilitarian "movement from one spatial unit to another."29

This was a strikingly sweeping argument to make—and for many of the reviewers, Baker’s attempt to grapple with it was mixed, at best.30 One reviewer lambasted the "Brobdignagian" tome for maintaining a "ratio of one interesting idea to about 200 pages."31 Another, ornithologist and SUNY Albany professor Kenneth Able, panned Baker's oeuvre in a 1979 review, calling his approach "too biased for the specialist, while too technical and not sufficiently balanced for the layman."32 Perhaps the only unequivocally positive review came in British science magazine New Scientist, courtesy of University of Liverpool biologist Geoffrey
Parker. Of course, Parker also knew Baker well: They had been classmates together at the University of Bristol, under Hinton’s tutelage, and had collaborated together.

Even though the book failed to garner accolades, it was an enormous steppingstone for Baker. When researching the book, he first stumbled across the massive amount of work done in homing behavior—and the curious absence of Man among the mix. And then it clicked. “Suddenly it made sense (to me at least) to study animal navigation using humans,” Baker said to me in an October 15 email, given how much easier they are “to deal with and to get to point towards home.”

The week after Parker praised his book in the pages of New Scientist, Baker followed up, summarizing it in a 1978 feature article in New Scientist the very next week. After working through his main arguments, he ended his article on an intriguing note, one that would echo through the halls of behavioral science for the next decade. “Experiments on human navigation that I am carrying out in and around Manchester show that...even modern industrialist humans make use of compass and landmark information in building up their familiar area.”

Exorcising Mesmer’s Ghost

As New Scientist readers leaned back and pondered Baker’s teasing conclusion, biologists had only just begun to show that animals sensed magnetic fields, opening up a whole new imagining of the “compass sense.” The field was vibrant, young, and exciting. Until about a decade before, the mere idea had been long dismissed as the den of charlatans and cranks. Scientists had long been skeptical of any claim that life could detect magnetic fields, especially the weak ones that ring the earth, a byproduct of our planet’s metallic core. Not only was there a lack of unambiguous physical evidence; there hadn’t been a good explanation for how life could sense magnetism. Where could there be a biological “compass,” what could it be made out of, and how would it work?

It didn’t help that early forays into the magnetic sense were deeply flawed. In the early 1770s, an Austrian physician named Franz Mesmer gained instant notoriety for his idea of a “universal fluid” in all living things, which he claimed could be manipulated with magnets and massages to cure diseases. Though he “mesmerized” wide swaths of the public, Mesmer’s claims were a stretch for many. He was quickly chased out of Vienna to France, where he continued trumpeting “animal magnetism,” much to the consternation of local academics. But Mesmer soon gained a protégé in Charles d’Eslon, a faculty member at Paris’ leading medical school, forcing King Louis XVI to assemble a commission in March 1784 to assess Mesmer’s claims.

From April to August, the commission, featuring chemistry juggernaut Antoine Lavoisier and American ambassador Benjamin Franklin, tested Mesmer’s “cures” extensively. This commission, with d’Eslon in tow, even traveled to Benjamin Franklin’s house in the Parisian suburb of Passy, where d’Eslon was asked to treat patients in front of witnesses. Over two days, fourteen patients of all ages—including Franklin himself—were treated with Mesmer’s techniques, but nine of them exhibited no improvement, and only three showed any kind of getting better. And in these three, as Lavoisier noted later, “the imagination alone, struck or primed to a certain point, [was] sufficient to produce these [salutary] effects.” The commission’s
findings were damning: There wasn’t anything “magnetic” about Mesmer’s claims. Instead, it was what would later be called hypnosis, an elaborate trick of the mind. Mesmer left France in disgrace, leaving the scientific community vindicated and a curious public smarting from a high profile debunking.  

Stained by Mesmer’s quackery, magnetic sensing remained out of vogue. The German naturalist Alexander Theodor von Middendorff hypothesized in 1855 that Siberian migratory birds had some kind of magnetic sense, but he was mostly shouted down, his ideas ungrounded in experiment or theory. The few experiments that were done hardly helped Middendorff and others’ case. In an October 1883 lecture, prominent British scientist Sir William Thomson described a colleague who had stuck his head inside a massive electromagnet. “What was the result of the experiment?” Thomson asked the crowd. “If I were to say nothing! I should do it scant justice.” His colleague hadn’t felt a thing. Thomson concluded that only “exceedingly powerful” magnets would be able to set off any magnetic sense, and only then, such a sense was “just possible.”

By the mid-twentieth century, however, the experimental tables had turned. Behavioral researchers were amassing increasingly convincing evidence that animals changed their behavior when the surrounding magnetic fields also changed. The most dramatic results came in the late 1960s, when German researcher Wolfgang Wiltschko began studying the European robin’s navigation. Normally, the captive birds would hover near the northern portion of their cages, indicating that the birds’ “instinct” told them to go north. However, when Wiltschko and his wife Roswitha shifted the magnetic fields around the cages, the birds’ orientation also shifted, tracking with the new magnetic north. Many researchers were deeply skeptical at first, but the Wiltschkos’ experiments, while difficult to run, were eventually replicated—opening up a brave new world of the magnetic sense.

Even though the behavioral evidence was pouring in, many scientists—echoing Thomson—harbored deep doubts about the existence of a magnetic sense, mostly on the grounds that there wasn’t a clear mechanism by which organisms could detect magnetic fields. For one, Earth’s magnetic field is incredibly weak—so weak at any given point on Earth’s surface, in fact, that an everyday horseshoe magnet is hundreds of times stronger. How could animals sense this utterly faint cue? The most straightforward way would be for organisms to contain chunks of magnetite, the magnetic form of iron oxide used in compass needles. But at the time, this was considered was impossible. Manufacturing magnetite in industry required temperatures and pressures far beyond those in living organisms.

In 1962, however, paleoecologist Heinz Lowenstam found that marine invertebrates called chitons used magnetite in their teeth as a hardening agent, demonstrating for the first time that living creatures could synthesize it. At first, people were merely astounded that the material could even be made in a living creature. But in 1975, discoveries of magnetite crystals in aquatic bacteria begged the question: what was this magnetite for? Now armed with hard evidence, people revisited the ideas von Middendorff had teased and found some surprising answers. Perhaps these small bits of magnetite could twist and turn along with changing magnetic fields, passing along signals about the magnetic field’s north-south orientation—and, with it, some useful information to an organism trying to navigate its environment?
In the next few years, other researchers would step up to the plate, showing that other organisms beyond bacteria showed similar abilities to orient to magnetic fields. Bill Keeton of Cornell University found that affixing magnets to the backs of homing pigeons’ heads disrupted their flight they traveled to and from their home ranges. Soon thereafter, a trio led by SUNY Albany professor Charles Walcott announced that he had found magnetite deposits in pigeon skulls—offering a tantalizing, tentative link between effect and cause.

A new paradigm had emerged: animals really did have a magnetic sense after all, and there seemed to be a plausible mechanism. It seemed reasonable enough to revisit the question if humans really had a navigational “sixth sense”—perhaps, even, one guided by magnetism. After all, why should it be that humans are the only animals blind to magnetic fields? And so Baker made his fateful entry into the study of the human compass sense.

**Tracking Down Humans’ Sixth Sense**

From 1976 to 1978, Baker spent the fall running his first navigation trials with humans, interested in an initially non-magnetic question: Like other animals, could humans “home” and point back toward a location from which they’d been displaced? Each experimental run began largely the same way, with Baker leading his student volunteers—all of whom had lived in Manchester for at least two years—to a nearby parking lot, where he blindfolded and crammed them into the back of a white Sherpa van. Asking them to remain silent, Baker would then drive them 8 kilometers to the south-southeast before going to one of seven “release sites,” which were anywhere from 6 to 52 kilometers away and offered no view of the Manchester city skyline. Students got nauseous on the ride—or worse. “We had our fair share of travel sick people,” Baker recalled to me, and “there was always somebody who had drunk too much before the trip and who had to get out of the bus blindfolded and relieve themselves.”

Baker’s subjects were then taken out of the Sherpa one at a time and asked to state the compass direction of the release site in relation to the University of Manchester. They were then asked to point toward the University, and Baker would measure the compass direction of their outstretched arm to the nearest degree. He then asked the students to remove their blindfolds, take two minutes to assess their location, and then point again toward the University, this time describing the cues they were using—or thought they were using—to decide where the University was.

But even though he hadn’t yet thought to test if magnets were involved in any way. That stroke of inspiration would come from Yorkshire Television, a TV network in northern England. Intrigued by the recent spate of magnetic experiments in other animals—and hankering for good television—television executive David Bellamy approached Baker and asked him if he were interested in doing a replication of the Wiltschko’s work, but in humans. Intrigued but hesitant, Baker vacillated, unsure if he wanted to perform this kind of experiment in front of television cameras. But his graduate student Janice Mather insisted that he do it, confidently asserting that the experiment would yield positive results. Fresh wind in his sails, Baker eventually agreed to conduct the experiment.
Presumably to keep the experiments closer to home for its Yorkshire audience, the network decided to conduct them in Barnard Castle, a small town some two hours to the north of Manchester. By design, Baker didn’t know when he would be conducting the experiment until June 29, 1979, the day of. On that drizzly Friday, he was given 42 local secondary school students as subjects, with two experiments on the docket for the day—one just like his previous one, and one with some extra “bizarre touches.”

The setup for the day’s first experiment was much like the ones Baker had run previously. Students were blindfolded and placed in an automobile, in this case a chartered bus that must have felt swanky in comparison to the Sherpa. Baker gave the students cards and writing implements, which they used to record their estimated compass direction from their school while still blindfolded and on the bus. After weaving through the Barnard Castle town center and taking a roundabout drive to a location some 22 km north of their school, the bus came to a stop, allowing the still-blindfolded students to record on their cards where they thought they were in relation to their school.

But in a special wrinkle suggested by Yorkshire Television, Baker ran a second set of experiments with 31 of the schoolchildren. Like the other experiments, the blindfolded students were loaded onto the bus and carted about, this time driven to the southwest and then southeast of Barnard Castle to two different stops. At each, the students were asked to record what they thought their compass bearing was in relation to their school. This time, though, the students had metal rods tucked into their blindfolds’ elastic, which the students were told were magnets.

In reality, half were magnets and half were inert brass bars, allowing Baker to see if magnets really did have an impact on this apparent “homing” behavior. And in his subsequent analysis, Baker broke up these test subjects and the Manchester participants into as many different categories as he could think of, just to see if there was any impact. How did the students taken only 7 kilometers away fare in comparison to those taken 52 kilometers away? How did students do in cloudy versus sunny conditions? How did students taken to the southwest do, in comparison to those taken to the northeast? And what happened once magnets got involved?

Any way he sliced it, Baker’s results seemed conclusive, and stunningly so. Better than chance, his Manchester subjects seemed to be able to orient themselves back to the University of Manchester, regardless of the weather conditions, familiarity with the area, and the distance they were taken away from “home.” His work with the Barnard Castle students was equally conclusive. Overwhelmingly, the students taken north of Barnard Castle were able to indicate the direction of Barnard Castle, all while blindfolded and kept on board a bus.

Most surprising of all, magnets seemed to disrupt this “homing” sense. The 15 students with magnets strapped to their heads weren’t able to orient toward Barnard Castle, while the 16 students wearing brass bars had no problem doing so. Baker was enthralled; he’d stumbled on something that was potentially huge. Hardly missing a beat, he immediately drafted two papers on the results and fired them off to the prestigious British science journal Nature, which promptly rejected them. Undeterred, he sent a draft of his work to Bill Keeton and kept refining his second paper in hopes of submitting it to Science, Nature’s longtime American rival.
Amid the hubbub, Baker found time to conduct a third series of human homing experiments, this time back at Manchester. Over several weeks in October and November 1979, Baker revisited his second Barnard Castle experiment, this time investing in some magnetic headsets made of coiled wire. Half of the helmets were actively magnetic, powered by fully juiced battery packs, while the others were duds, their batteries intentionally drained. This experimental series yielded Baker’s best results yet. Not only did subjects fail to orient toward home when they wore the actively magnetic headsets, but depending on the polarity of the magnetic field, subjects with active headsets were disoriented differently.

Despite his compelling results, Baker failed to include them in his Science manuscript. Perhaps this was for the sake of time and space. However, Baker may also have been intending to leverage the new results in a separate publication: a new book, the deal freshly inked, about “human navigation and the sixth sense.” Once again, Baker would tango with publishers Hodder and Stoughton, who were clearly keen to capitalize on Baker’s pending fame. They certainly didn’t make it easy on Baker: Over a feverish four weeks in February and March 1980, Baker pounded away at the manuscript, writing 18 hours a day, seven days a week. Three things pushed Baker on: faith in his results, an “endless” supply of coffee, and the music of Mozart and Blondie, which blared as he pored over his work.54

With his book manuscript complete, Baker doubled back to finishing the paper that ostensibly justified it, submitting a first draft to Science in early June.55 That same month, migration researchers from across the American northeast descended on Rockefeller University’s field lab in Millbrook, New York, for a small, weekend-long symposium on migration and orientation.56 As presentations came and went throughout, whispers of Baker’s stunning results and pending publication spread like wildfire among the attendees. Had Baker actually found a magnetic sense in humans? By the time the symposium ended, at least three labs—one at Cornell, one at Princeton, and one at SUNY Albany, led by former Baker critic Kenneth Able—walked away avowing to try replicating Baker’s results.57

Word got back to Baker, who felt pleased and empowered, pricking his sides as he wrote more about his yet-unpublished research, this time the cover story for the September 18, 1980, issue of New Scientist. The cover illustration embraced Baker’s findings: A blindfolded pigeon, wearing a suit, calmly points his wing as if sure of where to go. The bird’s other wing-hand clutches a detector, wired to a headset topped with a cartoonish horseshoe magnet. The drawing’s subtitle is unapologetically bold: “Compass in the brain – a new human sense.”58 In the article, Baker had no problem with jettisoning any pretense of uncertainty:

Speculations concerning a “sixth” sense in humans have been rife for just about a century, and during that time, a magnetic sense has always been one of the main contenders...The Manchester and Barnard Castle experiments have shown not only that the sense exists but also that apparently we can all make use of it.59

Baker, however, still had edits to complete on his paper, which Science received on October 1 following suggestions from the article’s reviewers.69 But this was of little consequence to the earnestly confident Baker. Paralleling the coquettish conclusion of his 1978 New Scientist article, Baker once again teased the progress of his research, this time indicating that “word of the
Manchester discovery has spread through the scientific grapevine,” revealing that “at least six laboratories around the world [were] actively engaged on research into the human magnetic sense.”

Going in the Wrong Direction

At Princeton University, biology professor James Gould was surprised by the news of Baker’s work. He had spent years studying the intricate behaviors of honeybees, demonstrating conclusively—in an experiment he designed as an undergraduate—that bees communicated by “dancing” once they returned to the hive.61 He also knew a thing or two about magnetoreception. Not long before Baker’s paper made the rounds, he had discovered magnetite deposits in honeybees’ abdomens. As the bees’ small nuggets of magnetite strained to align with the earth’s magnetic field, it seemed as though this gave the bee a sense of where north and south pointed—potentially helping them get to and from their hives.62 It was a major find, but an experimentally annoying one to run. Raising bees and then gauging their behavior isn’t exactly easy work.

For Gould, the big takeaway from Baker’s purported results was that Gould could use humans instead of bees in the experiments he had been running. The change would be nothing short of revolutionary for the science. What could be better than having your experimental subjects actually point toward where they were inclined to go? “So then here comes this paper by Robin Baker,” Gould recalls, “and we thought, ‘By God! How could we have been so stupid! It’s so easy! Humans can do it!’” And using humans also opened up a world of large sample sizes. “You just promised your class a picnic, and there you had 120 subjects,” says Gould today, his voice dripping with sarcasm. “It seemed like a miracle. How could we be so silly not to try it before!”63

Eager to build on Baker’s results, Gould immediately set about to replicate them. But the more that Gould pored over the drafts of Baker’s Science paper, the more that he saw its potential pitfalls. For one, Baker’s blindfolds were everyday sleep blindfolds and appeared to be too thin for reliable use during the day, when he had run his experiments. And because the sleep blindfolds weren’t large or fitted enough to cover the bridges of people’s noses, Gould surmised, subjects could look straight down and catch glimpses of the ground underneath them—enough, perhaps, to affect the participant’s ability to orient.

Gould realized he needed a truly foolproof blindfold, so he turned to an expert—James “The Amazing” Randi, a local celebrity and magician famous for debunking charlatans claiming paranormal powers.64 Randi “knew something about defeating sensory-deprivation devices,” Gould later wrote. Randi once performed an illusion in which he drove through an obstacle-filled parking lot blindfolded, impressing Gould enough to give Randi a call.65 To Gould’s delight, Randi agreed to help out, designing a loose, double-layered blindfold that seemed to block all incoming light. And just to be sure that light wasn’t affecting the subjects, Gould used black felt hoods to cover the subjects’ heads in some trials. One can only imagine how creepy Gould’s setup looked from a distance.

But the blindfolds weren’t the only possible problems. The vans and buses that Baker had used to ferry around his students had enormous windows, but nothing in his paper suggested that
the windows were sufficiently covered. This opened up even more room for throwing the experiment’s results. Perhaps participants in the experiment could feel the heat of the sun on their faces—or maybe they could even guess where the sun was through the blindfolds—so Gould resolved to cover his buses’ windows with aluminum foil.

In the late fall of 1980, Gould set out on his first round of replications, testing 40 Princeton undergraduates using his revised method. Driving on a topsy-turvy route, Gould took his subjects about 20 kilometers west of Princeton’s campus. Half of the students had magnets tucked into their blindfolds, while half had metal bars, just like Baker’s Barnard Castle setup. But Gould’s first results were less than promising. On average, the students without magnets on their heads pointed over 90 degrees away from where they should have been pointing, and even then, their clumping was essentially noise in the data. Strangely, the students wearing magnets actually performed slightly better than those without the magnets, though they were still considerably off the mark.

Gould continued, this time trying to replicate Baker’s second round of Manchester experiments. He gathered up 15 graduate students and staff members, blindfolded and hooded them, and transported them on a circular route taking them between 8 and 18 kilometers away from Princeton. At each stop, Gould asked the subjects to write down their estimates for where Princeton was. The results were largely a wash. At five of the six stops, the subjects didn’t point back to campus in any appreciable way. At one of the stops, though, the subjects pointed in largely the same way—but away from Princeton, exactly the opposite of what they had been asked to do. He repeated this experiment on two different occasions, with the same disappointing results. Unlike Baker’s students, Gould’s just didn’t seem to be able to point home.

Meanwhile, Kenneth Able began attempting replications of his own back in Albany. In concert with his research assistant William Gergits, Able orchestrated experiments similar to Gould’s, but this time, he performed them at night, in an effort to minimize the influence of the sun. Able and Gergits’ results were similarly troubling. Subjects seemed to orient randomly, missing SUNY Albany by an average of 80 degrees in their estimates. Similar results accrued in Ithaca, New York, as Bill Keeton—tantalized by an initially successful trail—soon found himself mired in similarly fruitless results. Tragically, Keeton’s efforts to study the human magnetic sense were cut short. He died suddenly in August, casting a pall on Gould and Baker, who would dedicate his book Human Navigation and the Sixth Sense to him.

Finally, on a day freighted with the mystic—Halloween 1980—Baker’s article on humans’ magnetic homing finally hit the press. He had done it, showing “goal orientation by blindfolded humans after long-distance displacement” and the “possible involvement of a magnetic sense.” Media outlets immediately lavished Baker’s paper with praise. After all, who could resist the allure of an actual sixth sense?

First out of the gate, Phil Hilts of The Washington Post effusively covered Baker’s paper a week before its official publication, claiming that “our heads are natural magnetic compasses...that can tell us the rough direction of home even if we have been carted 40 miles away, on roundabout routes, blindfolded.” The New York Times followed suit less than two weeks later, echoing the refrain that “many of [the subjects] were able to aim themselves in the
right general direction," a homing ability that magnets "seemed to upset." The Times pointed to "reports in Science and the New Scientist" as sources—citing only articles authored by Baker himself, hardly batting a critical eye. Gould's fears had come true. The magnet-crazy media had taken the bait.

In the aftermath of the paper's publication, Baker felt that he had broken through. Now, finally, he had made his mark on the scientific enterprise, contributing something of immense value. "I was 36," Baker wrote to me, "and really thought I'd discovered something important." But the clouds around Baker were already darkening. Less than two weeks later, Gould submitted a first draft of a paper to Science rebutting Baker's claims.

***

Meanwhile, Tom Dayton—soon to enter the fray—was waiting to get his scientific career off the ground. At the time, he was almost a college graduate, awaiting his diploma from the New College of Florida for his bachelor's degree in experimental psychology. As he floated in academic limbo, he was stuck making ends meet, working the graveyard shift at a local yogurt yeast factory as a janitor and unofficial computer technician. His undergraduate thesis, a pithy examination of meditators' brain activity, had a certain ring to it—so why not submit it as a paper to Science? He knew that he had gone out on a limb, and Science quickly reminded him of it by promptly rejecting his submission.

But as Dayton flipped through the Halloween 1980 issue of Science, he came across Baker's article and was immediately struck by the issues the paper seemed to have. "We experimental psychologists spend huge amounts of effort to avoid biases," says Dayton, now an interaction designer with appointments at UC Santa Cruz and NASA. From an experimental psychologist's perspective—even that of a college student—Baker had taken some missteps, failing in Dayton's eyes to control for every possible influence that might have steered subjects astray. "No way did he adequately control his experiment," Dayton recalls thinking to himself. "Here's what I thought was a really crap article. How'd that get published, for God's sake?"

A couple of months later, Dayton was reminded of Baker's follies once again—this time, by magazine articles in SCIENCE 81 and Psychology Today by James Gould, who had taken to Baker's seemingly favorite medium. Taking a page out of Baker's playbook, he had gone public with his initial results before publishing them officially, expressing doubt over Baker's reported results in an inset to Baker's latest triumphant magazine article. Excited that someone else had trepidations, Dayton wrote to Gould. Much to Dayton's surprise, Gould responded, inviting him up to Princeton to help with a scheduled replication of Baker's study—one that Baker would be attending. Stunned by Gould's reply, Dayton wasn't sure if he should go, but his wife encouraged him to do it. And so off he went, entering what he'd later recall to me as a "battleground in experimental psychology."

On February 15, 1981, Baker joined Gould, Dayton, and Joe Kirschvink in Princeton. While the three-day visit began perfectly pleasant—everyone noted that Baker was personable, confident, and ready to work—tensions began to mount on both sides. Baker recognized almost harsh levels of skepticism in Gould, while Gould became increasingly exasperated. As it
turned out, bits and pieces of Baker's procedure hadn’t made his Science paper. Instead, they were tucked into Baker’s book or sprung on Gould during the visit—which didn’t exactly sit well with him.80

First, it turned out that Baker took the subjects in his original Manchester experiments to the Zoology Department roof before driving them away from campus, allowing them a bird's eye view of the local topography and the distant horizon before getting in the van. Baker also acknowledged that in some trials, sunlight hadn’t been controlled for as well as it could have. The white Sherpa van’s windows were inconsistently covered, allowing sunlight into the back of the van.

The Barnard Castle experiments also had unmentioned problems. Baker had asked the students get out of the bus at each stop and point toward the school both with and without blindfolds. However, concessions had been made for telly. Yorkshire Television insisted that all the students be taken out of the bus at once, leading to precisely the raucous mess one would expect. Baker, to be fair, wasn’t thrilled with this, only including the written estimates from the Barnard Castle experiments. Problematically, however, he never acknowledged that these trials had failed—nor did he attempt to see how they could have affected the tests’ written portions.81

Baker and Mather left Princeton after three days of mounting tension and round after round of negative results.82 Kenneth Able and William Gergits, working in Albany, reached similar conclusions. They hadn’t been able to replicate Baker's results, either. At the last minute, Able and Gould joined forces—knocking Gergits out of a Science co-authorship, to his everlasting chagrin83 —and had their combined publication hit in late May 1981. While the paper pretended to be conciliatory in a tacked-on note, the subtext could not be clearer. Baker seemed to be mistaken.84

Chastened somewhat but unbowed, Baker countered. Perhaps the American students’ polyester clothing ensured they picked up more static charge, throwing off the results?85,86

Though his static explanation would win him few fans, Baker was soon dealt a nasty shock of his own: an increasingly bemused and skeptical media. In particular, Paul Jacobs of the Los Angeles Times took Baker to task. After savagely reviewing his 1981 book, Jacobs penned what was then the definitive American news coverage of the Baker saga.87 It was none too flattering. “To get Baker’s results, you’d have to be a really sloppy scientist,” Kirschvink said to Jacobs—a denunciation that still perturbs Baker, over three decades later.88,89

In fact, outside of the 1982 Naturewatch program I had first watched, “which gave [his] work a welcome and totally serious treatment,” Baker emphasized, the media had started lumping him in with the very fringe he thought researchers of his ilk had escaped. In one email he sent me, Baker recounted a cringe-worthy appearance he made on a German television show in 1984, where he was unwittingly put on a panel filled with “dowsers [and] telepathists.”

“It turned out that their whole intent was to make myself...look stupid,” he wrote, “and they succeeded.” Without decent in-studio translation—Baker doesn’t speak German—he was left floundering, “looking a total fool in the question-and-answer session.”
"I should have walked out once I realised what was happening – except that I thought I could handle the situation," Baker wrote to me. "And by the time I realised I couldn't, it was too late."

Smarting from popular and scientific criticism, Baker aimed to silence his critics once and for all, mounting his counterattack. In a 1983 *Nature* paper, he claimed that he had located humans' magnetic sensory organ. It seemed to be exactly where Baker thought it would be: deep in the sinus, taking the form of a small bone containing elevated levels of magnetite. Intrigued, Kirschvink—now a professor at Caltech—asked for samples of the magnetite, and Baker obliged. But within hours of receiving the samples, Kirschvink determined that the samples were contaminants, artifacts from the surgical steel saw that Baker had used to slice open the bone. Gould, now thoroughly annoyed with Baker and the "waste of time" that his research has become, welcomed the news from his former Princeton colleague.

Baker eventually admitted that his magnetite samples were contaminated, but he kept pushing his ideas further, still convinced that humans had a magnetic sense. He adopted better blindfolds and developed two new versions of his magnetoreception experiments, most notably a new "chair" experiment. Blindfolded subjects were spun around in a chair and required to guess the directions of the compass rose dozens of times in a row. For years, Baker's graduate student Gai Murphy would spin the chairs, culminating in her 1987 PhD thesis on "the development of compass orientation in children."

Reading it over, it's clear whose team she was batting for. She cites Baker, her adviser, amply and approvingly, using him to buttress her argument that the hundreds of children she tested could orient toward given compass directions better than chance, a sense that apparently develops during adolescence. However, her treatment of Gould and Able's 1981 paper—or Kirschvink's 1985 failure to replicate the chair experiments’ results—is considerably briefer, dismissing them and other negative results as either procedurally flawed or actually supportive of Baker's claims. It's a bold statement, one that would have been interesting to discuss with Murphy herself. But for whatever reason, she now doesn't want to touch her dissertation with a ten-foot pole. Murphy repeatedly refused to comment when I contacted her office at England's University of Central Lancashire, where she was recently appointed Pro Vice-Chancellor. Her assistant says that Murphy hasn't spoken to Baker in over twenty years.

It must have been interesting for Murphy to flip through *Magnetite Biomineralization and Magnetoreception in Organisms*, the book where Kirschvink published his results. While most of the tome is stately and staid—ostensibly, it's a vast review on magnetoreception in general—it eventually devolves into an all-hands-on-deck fight over Baker's work. Baker himself delivered the first punch in Chapter 26, asserting that Gould and Able's 1981 paper contained incorrect—and misleadingly negative—data and results. Up next was the young, scrappy Tom Dayton, who claimed that Baker chose imperfect statistical tools several pages before. The Albany crew—Kenneth Able and a finally published William Gergits—followed, but little good news to offer: Except for a single, inexplicably positive trial, Able and Gergits found no evidence of magnetic human homing.
Chapter 29, arguably the most damning, came from Kraig Adler and Chris Pelkie, the duo at Cornell who had taken up Keeton’s mantle. They, too, had invited Baker to help with replications, and he obliged on February 20, 1981, days after his fateful Princeton visit. He ran four trials in Ithaca using his personal method, and lo and behold, he got significant results: His subjects appeared to have a magnetic homing sense. But then Adler and Pelkie took systematic steps to remove the influence of Baker when they re-ran the experiments the following year. Instead of having someone instruct subjects in the flesh—where a flick of the wrist or a slight change in voice could accidentally influence a subject’s decision—they instead recorded instructions on tapes and signs. And once they did, they stopped getting signs of homing.

They also took their inquiry one step further. Adler and Pelkie generated random sets of data that mimicked the kinds of numbers that Baker had been getting. But Adler and Pelkie’s artificial data sets were intentionally skewed. They suspected that Baker was unknowingly preventing subjects from pointing in certain directions. To see how sensitive Baker’s statistics were to this sort of accidental meddling, they made data sets that excluded increasingly large swaths of 360 degrees. Some were made as if participants couldn’t choose a five-degree chunk. In others, this excluded chunk was ten, twenty, thirty degrees, sometimes even more. How much did you have to prevent people from choosing, Adler and Pelkie wondered, to get the false impression that the increasingly confined data points were “pointing” somewhere?

The results were stunning. Of the 360 possible degrees in which subjects could point, Adler and Pelkie showed that if biases only cut out 20 degrees—that is, a subject could choose 340 of 360 degrees—then that could be enough to throw the stats. Slice the pizza of possibility into eighteen slices, remove one, and you could get a false positive. In effect, Baker’s results could be real—or they could be meaningless, bled to irrelevance by a thousand pinpricks of bias.

As if to close, James Gould once again entered the fray, a grumbling master of ceremonies over the whole critique of Baker’s methods. In a terse five pages written by someone with strong opinions but clearly sick of debating them, Gould laid out his views starkly:

To summarize, Baker’s calculations appear to be incorrect and potentially misleading; his techniques are sometimes questionable; and his results cannot be replicated. I conclude, therefore, that humans probably lack any significant ability to sense direction magnetically.

Needless to say, it’s difficult to put down the book and say that Baker came out of it a winner.

Outside of the 1985 debate-in-a-book—capped off by an earnest rebuttal from Baker in Chapter 34—Baker’s assertions still ran into issues. Replication attempts at the University of Tulsa and Monash University came up with nothing. Student groups, most notably at the University of Sheffield and Swarthmore College, gave Baker’s tests a go, to mixed results at best. And no one was able to demonstrate the magnetic sense with the precision and convincing directionality that Baker was able to achieve. Baker, though, was convinced that actual positive results were hidden in these authors’ data, an argument he lobbed into the fray in a 1987 study and 1989 book. However, these did little to change anyone’s mind. Baker was being left behind.
Soon enough, Baker would return the favor, finally leaving animal navigation altogether to pursue research on human sexuality.\textsuperscript{104} Why did he leave? Was he fed up with the ongoing skepticism, as many suggested? Starved of new research funding? Baker certainly was fed up with what he saw as unfair criticism—and there were certainly whispers that Baker's funding was on shaky footing\textsuperscript{105}—but according to Baker himself, he was merely moving to greener pastures. "By 1986, there was nothing more I could do in the field apart from actually identify the sense organ itself, which I suspected...could have taken the rest of my working life," he wrote to me. "By then, anyway, I'd had a couple of...ideas about human sperm competition, found an equally excited co-worker, and really didn't want to be distracted by anything else.

"So I collected together all my unpublished magnetoreception work and conclusions, put them into a book so that they would at least be available to future workers – and left the field.

"I wasn't driven out and I wasn't disillusioned—it was simply time to move on."

\textit{The Science of Things That Aren't So}

On December 18, 1953, a group of scientists ducked in from the snowy winds outside of New York's Knolls Research Laboratory. They were on their way to a talk, and from a big name, too: Irving Langmuir, a Nobel Prize-winning chemist who represented the best of what science was. Langmuir's talk, however, would become famous for its focus on science's ugly temptations—what he called "pathological science." Looking across the landscape of science, Langmuir identified example after example of cases where honest researchers fell "into false results by a lack of understanding about what human beings can do to themselves"—a trap of self-deception in which Robin Baker may have later ensnared himself.

Langmuir's first example of this pathology was one of the most notorious scientific flubs of the early 20th century—an embarrassing fiasco from the field of physics. In 1903, respected French scientist Rene Blondlot claimed in a report to the French Academy of Sciences that he had discovered a companion to X-rays, a bizarre form of radiation he called "N-rays." Produced by heating up a wire behind an aluminum plate, they were only detectable when projected through the aluminum onto nearby surfaces in almost complete darkness. All sorts of items gave off N-rays—humans, bricks, the Sun—but the rays weren't visible unless it was absolutely silent and sufficiently warm in the room in which they were to be observed.\textsuperscript{106} In the coming months, a slew papers were published on N-rays, and about half of them confirmed Blondlot's results. Perhaps there was something to these N-rays after all?

Intrigued, a scientist named Robert Wood visited Blondlot's lab in 1904 to check things out for himself. Blondlot eagerly showed off a new experimental setup: He had fashioned a prism out of aluminum, claiming that he could split N-rays into different components, like glass prisms separate visible light into colors. After examining the device, Wood was confused. Blondlot claimed he was getting measurements better than his equipment seemed to allow. Intrigued, Wood asked him to demonstrate how the prism split up the N-rays. Blondlot happily obliged, fastidiously working until announcing results similar to other trials. There was one catch, though. Wood had snuck the prism into his pocket before asking Blondlot to repeat the measurements. In a brutally terse article in \textit{Nature} published soon after his visit, Wood concluded: "[The] few
experimenters who have obtained positive results have been in some way deluded.” The N-rays were all in Blondlot’s head.\textsuperscript{107}

In recounting this story to his audience, Langmuir wasn’t saying that Blondlot was malevolent in his fraudulence. On the contrary, Blondlot was sincere in his belief that he had discovered N-rays. But dangerously, Langmuir emphasized, Blondlot had tricked himself into thinking that N-rays existed. And Blondlot’s N-ray scandal is but a link in a long chain of imaginary results. Perhaps subjective bias crept into experiments through unforeseen avenues. Perhaps the investigators were blissfully infected with wishful thinking. Or maybe researchers misunderstood the kind of responses they were seeing in their experiments. At any rate, these so-called “pathological” results—doomed by human biases—carried six symptoms, according to Langmuir:

1. The observed effect has a nearly undetectable, low-intensity cause.
2. The effect is small and barely measurable or requires enormous sample sizes.
3. Whoever finds the effect claims that their measurements are extremely accurate.
4. The theory behind the effect is out of left field.
5. The effect’s observers always have an answer for criticism—even if it’s off the top of their heads.
6. No more than half the researchers who ever study the effect manage to reproduce it.\textsuperscript{108}

Perhaps the best-known modern example of “pathological science” occurred in March 1989, when University of Utah chemists Martin Fleischmann and Stanley Pons announced at a flashy press conference that they had observed nuclear fusion at room temperature. They had electrified rods made of precious metals and submerged in water made with deuterium, a heavier version of hydrogen. Excitedly, they claimed that the water was heating up far more intensely than it should have been. This “excess heat” was so unexpectedly large, they announced, that it could only come from pairs of deuterons—isolated, charged deuterium atoms—fusing together inside the electrified metal rods.\textsuperscript{109} The news rocketed around the world. Here was a discovery that would forever make energy production clean and affordable!

However, the scientific community, led by physicist Dick Garwin, quickly tore Fleischmann and Pons’ results apart. Labs the world over immediately tried to repeat the tabletop experiment but saw no excess heat. Moreover, physicists were highly skeptical of the idea that atoms could fuse together at room temperature. Sixty years of physical theories said that fusion only happened at massive temperatures and pressures, in order to overcome the repulsion that two atoms’ positively charged nuclei have against one another.

What’s more, condensed matter physicists had a field day examining the metal rods. Fleischmann and Pons had claimed that loading deuterons into the metal rods’ palladium lattice would be like keeping the deuterons under incredible pressure, increasing the chance that they would collide into one another. But at the atomic level, the distance between any metal atom and the deuteron wedged next to it was far too large to account for Fleischmann and Pons’ proposed fusion by squeezing. Unless they had a theory that allowed for contradictions to decades’ worth of theory and evidence, Fleischmann and Pons were simply mistaken.\textsuperscript{110,111}
Incidentally, Dick Garwin—pivotal in the debunking of cold fusion—played a sizeable role in debunking another spurious scientific claim twenty years before, one involving a University of Maryland physicist named Joe Weber. Like many in physics, Weber was keen on proving the existence of gravitational waves, ripples in space-time generated by objects as they speed up or slow down. However, while gravitational waves were fairly well established in theory, they had never been detected directly. This is because the predicted waves are extraordinarily weak.

But Weber thought he had cracked a way to detect them. He isolated two massive aluminum cylinders in two separate vacuum chambers set some distance apart, each covered with sensors that emitted electricity if they were bent. Weber hoped that a passing gravitational wave would strain the first aluminum bar, continue wiggling through space, and then warp the second aluminum bar. In each, the sensors would convert these minute vibrations into electric impulses magnified by an amplifier. If a wave actually sped by, two successive blips would appear on his detector’s screen. Armed with his detector, Weber stood at the ready, a rookie highway cop eager to write his first ticket.

In the late 1960s, he claimed that he detected these minute waves with his device, a feat that detractors soon claimed was physically impossible and likely an instrumental error. After 1974, Weber would fall out of the scientific mainstream, convinced until his death in 2001 that he had observed gravitational waves. While Weber’s “acoustical cylinders” underlie most of the technology behind today’s gravitational wave detectors, his own recklessly generous interpretations of his data left him without a leg to stand on.

Of these classic examples of irreproducible results, Baker’s rise and fall most closely mirrors Weber’s. Despite the eventual skepticism surrounding their work, both had pushed forward in good faith, with existing theory supporting them. In fact, like Weber’s critics, everyone is in agreement that Baker’s initial research question was completely legitimate. And Baker never proposed any theories for human magnetoreception outside of those proposed by other researchers for other animals. And all of these theories were predicated on the well-documented presence of magnetite in animal tissues. He was merely seeing if the well-established phenomenon of magnetoreception also applied to humans.

But Baker’s earnestness and initial reasonableness don’t absolve him from legitimate criticism under Langmuir’s framework. In defending his stance, Baker pointed to the surprising—and unique—strength and accuracy of his results. What’s more, he tended to blend fact and opinion and published surprisingly few studies in rigorously peer-reviewed publications, largely opting instead to vouch for the magnetic “sixth sense” in his own books and magazine articles. He also had a nagging tendency not to control for all possible factors that could have influenced his results, a possible holdover from the fact that he had no prior experience designing experiments for human subjects.

The challenge Baker faced—indeed, the one all scientists face—was isolating the one effect in which he was interested amid the distracting background of the raucous natural world, along with the accidental stimuli that experimental setups introduce. In the mid 2000s, Tyson Platt—a psychology professor at Alabama State University—inadvertently came across one of
these “confounders” while in graduate school at Auburn University. Platt’s PhD dissertation, completed in 2007, consisted of a three-part effort to replicate Baker’s work, the most recent I’ve identified. His longest-term study brought Baker’s ideas to their most basic: Could you condition someone to tell the presence or absence of a magnetic field around their head?

To test this, Platt and his colleagues built a wooden chair with magnetic coils near the head, wired to two heavy-duty mechanical switches. One turned the coils on and off, altering the magnetic field. The other was a dummy, disconnected from the circuit. By flipping one of these two switches, Platt could see if a test subject really could tell if a magnetic field was present or absent. To Platt’s extreme surprise, one of his test subjects seemed to be surprisingly good at telling the difference between these two states. How was he doing this? Did he really have Baker’s magnetic sense?

The truth was even stranger. The subject claimed that he could hear the difference in the clicks the two switches made, which allowed him to keep track of whether or not the magnetic field was on. And when Platt checked this by turning off the power and just flipping both switches, the subject was right. He could tell the difference on the clicks alone.

The subject’s uncanny hearing set Platt on a surprisingly cartoonish effort to eliminate the effect of sound. To get the subject to stop hearing the difference, Platt—who moonlights as an audio engineer—ultimately had to mount the switches in wooden boxes, stick those boxes in a soundproofed cooler across the room, give the subject heavy-duty earplugs, play static over speakers in the room, and play recordings of the switches’ clicks at random intervals. It’s worth mentioning that Baker—in fact, most scientists experimenting on human perception—didn’t take these extreme steps to eliminate acoustic biases.

These kinds of complicating factors harpoon researchers’ results all the time, and modern researchers largely conclude that Baker too had fallen victim. His protocols weren’t stringent enough to eliminate other possible explanations for his subjects’ ability to orient. Some could feel the sun’s heat on their cheeks. Others could see through his blindfolds. Others still could unconsciously incorporate subtle hints—like the orientation of the bus—in their responses. Despite his best efforts, he had left too much to outside factors and chance.

But Baker has stuck to his guns until the bitter end, swearing to this day that his results were real. “I’m not quite sure how to say this without sounding irritated,” he wrote to me, “but [at] Manchester, I and my colleagues spent ten years doing hundreds of rigorous experiments using thousands and thousands of different people under all sorts of different conditions. “We not only demonstrated the phenomenon was real but also worked out the effects of time of day, age, gender and how magnetoreception combined with sun and star orientation to produce navigational solutions.” People like Gould, Kirschvink, and others, however, “each spent a couple of hours doing one or two casual experiments with usually just a few people.”

In remaining defiant, Baker is waiting for his own vindication, too—regardless of whether or not it comes. In fact, at first glance, it might seem like Baker was onto something. A 1991 study claims to demonstrate that when trained for days, humans can feebly orient northward on command after being spun around. In 1992, none other than Joe Kirschvink found that there
were magnetite deposits in the human brain.\textsuperscript{116} A 2003 study suggested that changing the magnetic field around a person’s head could change her eyes’ sensitivity to incoming light.\textsuperscript{117} Another study, conducted in 2011, implied that fruit flies’ magnetic sense—mediated partly by their version of a protein called cryptochrome—also works with the human version of the same protein.\textsuperscript{118,119}

But these tantalizing studies don’t offer as much for Baker fans as they might hope. First, none of them conclusively demonstrate that humans have a functional magnetic compass sense. It’s worth emphasizing this point again: For the last thirty years, no one outside of Baker’s employ has been able to replicate his striking results. The 1991 study only examined eight people.\textsuperscript{120} While it’s unclear why our brains contain magnetite—it might be involved in memory formation, but we don’t know for sure\textsuperscript{121}—there’s no evidence that it’s used to sense the earth’s magnetic field. Finally, the 2003 and 2011 studies focus on a pathway for sensing magnetic fields that doesn’t involve magnetite. Instead, this proposed mechanism is sparked by light hitting proteins in the back of the eye.\textsuperscript{122,123,124} Baker, however, claimed that his subjects could orient while blindfolded.

Most tragically of all, some of Baker’s sporadic work in non-human magnetoreception has more or less held up, suggesting that had he confined his efforts there, his body of work would have better withstood scrutiny. In a paper published in \textit{Nature} in 1981, Baker and Janice Mather showed that the humble woodmouse appeared to have a magnetic sense of direction. Not much happened with these results—Baker continued to pour his efforts into studying humans—but in April 2015, a paper in \textit{Scientific Reports} confirmed Baker’s basic observation in spirit. While Baker focused on woodmice’s navigation, the recent paper shows that they prefer to build their nests on sides of a given enclosure in line with magnetic north and south. Shift the orientation of the surrounding magnetic field, though, and the woodmice’s roosting preferences unambiguously shift in kind.\textsuperscript{125,126}

Since Baker was seemingly on the right track in mice, does this mean that time will be kind to his human trials? Probably not. Baker was explicit in his efforts to treat humans just like any other animal, and they are. But humans are incredibly clever animals, extremely sensitive to all sorts of cues that Baker arguably failed to fully appreciate and account for. There will always be an asterisk of doubt placed next to Baker’s human studies, even if someone else steps up and reproducibly demonstrates that humans have a magnetic sense.

As of now, though, nobody is eager to take up his mantle.

\textbf{Passion, Science’s Guiding Star}

In the face of irreproducible results—and his peers’ skepticism—Baker also seems to have fallen victim to a parental love of his own work, keeping him too committed to his own ideas in the face of uncomfortable evidence to the contrary. For all scientists—pathological or otherwise—their work is the end product of years’ worth of tests, countless hours of thinking, writing, and dreaming. It’s a lot to let go of at the first signs of contradictory evidence. And so people across science tend to double down on their claims, for good or for ill. Blondlot doggedly
clung to his ideas about N-rays. Fleischmann and Pons’ ragtag successors swear up and down that they have unlocked the secret to clean energy.\textsuperscript{127} And Baker has stuck by his magnetic sense.

But there’s more to this pride than a simple matter of ego. Scientists, as a rule, have to be passionate about what they do. It’s not the kind of job that people clock in and out of. It demands a manic, monastic persistence, something that can motivate people through the long nights in the lab, the stress of academia, and the danger that experiments years in the making will yield absolutely zilch. This passion, though, is a double-edged sword. As it energizes scientists, it also runs the risk of clouding their minds—and making the professional deeply personal. As the philosopher of science Michael Polanyi put it in his seminal work \textit{Personal Knowledge}:

...Passion seeks no personal possession. It sets out not to conquer, but to enrich the world. Yet such a move is also an attack. It raises a claim and makes a tremendous demand on other men; for it asks that its gift to humanity be accepted by all. In order to be satisfied, our intellectual passions must find response. ...We suffer when a vision of reality to which we have committed ourselves is contemptuously ignored by others. For a general unbelief imperils our own convictions by evoking an echo in us. Our vision must conquer or die.\textsuperscript{128}

But what sustains this exhausting, do-or-die passion? A brief dip into early twentieth-century astronomy could help answer this question. At the time, a debate raged over the size and structure of the universe. Some astronomers believed that spiral-shaped, cloud-like objects across the sky dubbed \textit{nebulae} were all clusters of stars and gas within the Milky Way, our own galaxy. None were as “mulishly wedded” to this idea as astronomer Harlow Shapley, who held that the universe consisted solely of the Milky Way, on the basis that it was much larger than astronomers previously believed.

In spirit, Shapley was right about the Milky Way’s size: It’s about 100,000 light-years across, far larger than many astronomers thought at the time. But astronomer Edwin Hubble and others eventually showed that the nebulae that Shapley had considered part of our galaxy were much too far away for that to be true. In fact, the Milky Way was but one island in a vast, expanding cosmic archipelago, filled with billions of galaxies. At first, Shapley fought this idea tooth and nail, yet he eventually came around to it, later “becoming its most boisterous promoter”—a more enthusiastic proponent of Hubble’s discoveries than Hubble himself.\textsuperscript{129} So how was he able to divorce himself from his former, fiercely held position?

For one, Shapley was a world-class scientist who clearly saw where the winds of evidence were blowing, even though it meant abandoning his once-cherished ideas. In fact, Shapley regretted not changing his mind sooner: Had he done so, he might have gotten more of the renown reserved for Hubble. But there’s a certain sense of \textit{liberation} inherent in the massive, expanding work laid out by Hubble and others. To Shapley, this could have meant that there was more Nature out there to try and understand. And he spent the rest of his career doing just that. The sandbox had suddenly gotten bigger, and Shapley may have relished the chance to play in it.

This may well be another reason why people like Robin Baker stick by their results, despite substantial evidence against them. In their own work, they claim to expand our
conception of what is possible, asking and answering sexy “what ifs”: Can you just use an acid bath to make stem cells? Can a simple electrode setup solve the world’s energy problems? Can humans sense direction magnetically?

If these questions are answered with “yes,” then our imaginations can’t help but be inflamed with wonder. If they hold up, discoveries like Baker’s render the universe more beautiful, more awe-inspiring, and more complex. And that’s deeply satisfying. The party line in science is that an experiment like Baker’s is informative, regardless of whether or not the results are positive. But which feels better: proving that humans have a magnetic sense, or disproving that humans are more than what they appear to be? Once researchers latch onto a bewitching idea and then pursue it—fueled by passion and curiosity—it’s hard to admit wrongness, even as the din of other scientists’ disapproval grows ever louder. On what grounds do myopic naysayers question their groundbreaking work?

The brutal truth of the matter, though, is that the scientific community does have the authority to call those shots—and does so on the basis of what’s discernible on the group level. Science doesn’t happen in a series of cloistered labs, with individuals submitting their work to the communal record without any kind of vetting. It’s a boisterous debate, a clash of egos, a battle of wills and intellect from which truth—or our best approximation of it—settles out.

All of this requires that scientific results exist outside of their proponents. Those piles of serene textbooks bequeathed to us by science? They’re manufactured through grinding up and combining the pulp of individual scientists who are all too human.

***

It’s December now, and Baker is trying to wrap up our email contact diplomatically. I offer to fly to Spain to see him, but he declines, saying that he’s not sure what I would get out of it. Besides, he says, he has to finish his latest book. But he also needs to step away from the memories I’ve dredged up by contacting him. “Anyway, I’m getting carried away,” he says in one of his final messages, as if telling himself to stop.

“If I’m not careful, all my old enthusiasm will come back.”
Endnotes

1 Baker, Robin. Email. 14 October 2014.
12 Zwelling et al. A survey on data reproducibility in cancer research provides insights into our limited ability to translate findings from the laboratory to the clinic. PLoS ONE. 2013.
17 Ibid.
19 Ibid.
20 Baker, Robin. Email. 15 October 2014.


A translation of von Middendorff’s hypothesis: “...The thought is tempting that the astonishingly steadfast migratory birds—despite wind and weather, in spite of night—are aware [of magnetic fields], and therefore accurately observe migratory orientation...As ships use a magnetic needle, these ‘sailors of the skies’ would use this internal magnetic feeling...”


Wiltshko, Wolfgang, and Wiltshko, Roswitha. Interview. 11 October 2014.

Gergits, William. Interview. 27 January 2015.


Stoddard, Phillip. Email. 4 February 2015.

Gergits, William. Interview. 27 January 2015.


Ibid. 844-846.


Phil Hilts was the former director of the Knight Science Fellowship at MIT and a former professor in the MIT Graduate Program in Science Writing.


Dayton, Tom. Email. 19 February 2015.

Dayton, Tom. Interview. 26 January 2015.


Dayton, Tom. Interview. 26 January 2015.


The magician James Randi also participated in the Princeton trials, later claiming he could partially intuit his orientation through the attraction between the magnet he thought he was wearing and the metallic wall of the bus in which he was riding. However, Randi was actually wearing an inert brass bar—a fact that amuses Baker to this day.

Gergits, William. Interview. 27 January 2015.


Baker, Robin. Email. 27 October 2014.

Baker, Robin. Email. 27 October 2014.


Kilmartin, Catherine Ann. Email. 10 March 2015.


Partridge, Karen. Email. 3 December 2014.

Williams, Tim. Email. 6 May 2015.


