Is Geomagnetic Sensitivity Real? Replication of the Walker-Bitterman Magnetic Conditioning Experiment in Honey Bees

JOSEPH L. KIRSCHVINK AND ATSUKO KOBAYASHI
Division of Geological and Planetary Sciences, The California Institute of Technology, Pasadena, California 91125

SYNOPSIS. Although the presence of geomagnetic sensitivity has been suspected for a long time in a variety of marine and terrestrial animals, many responses reported in the literature have been based on extensive statistical analysis of orientation results or rely on obscure behavioral activities (like cetacean strandings or honey bee waggle dances.) None of these reports have yet approached the level of clarity and simplicity displayed in experiments with the magnetotactic bacteria, which is the best example of geomagnetic sensitivity in any living organism. Furthermore, claims of magnetic effects on living organisms pervade the literature of biophysics, but many have failed subsequent attempts at replication. We need to develop simple and easily replicated experiments for marine and terrestrial animals which can be modified to answer basic questions concerning the psychophysics of any geomagnetic sensory system which might be present. In this paper, we report the first replication of the Walker-Bitterman magnetic anomaly conditioning experiment in honey bees, as well as one of our attempts to slightly alter their basic protocol. We also report our attempts to condition honey bees to magnetic direction in simple maze experiments, and the initial results of a pulse-remagnetization experiment designed to test the ferromagnetic transduction hypothesis. We conclude honey bees are sensitive to the geomagnetic field, that the signal processing for it is more complex than previously thought, and that a ferromagnetic transducer is compatible with all known behavioral data.

INTRODUCTION

Despite over a century of work by numerous people, the presence or absence of geomagnetic sensitivity in animals remains one of the most controversial problems in the behavioral and neural sciences. Two questions are of fundamental interest in this regard, the first being whether any of the higher animals are able to perceive magnetic stimuli as weak as that of the Earth's field. If so, the second problem concerns the mechanism by which they achieve this sensitivity. Earlier in this century the prevailing dogma held that the geomagnetic field was far too weak to produce measurable effects on living organisms (e.g., Griffin, 1944, 1952), and many claims of magnetic sensitivity in homing pigeons (e.g., Yeagley, 1947) were attacked on the basis that the induction-based mechanisms proposed for the sensory transduction were judged implausible, and no other mechanisms were known.

During the last two decades, the discovery of ultra-sensitive electrophoretic sensors in elasmobranch fish (Kalmijn, 1978) as well as the discovery of magnetite (Fe₃O₄) biomineralization in the teeth of polyplacophoran mollusks (Lowenstam, 1962) demolished this "implausibility" argument. Subsequent discoveries of biogenic magnetite in the magnetotactic bacteria (Blake more, 1975), honey bees (Gould et al., 1978), homing pigeons (Walcott et al., 1979) and pelagic fish (Walker et al., 1984, 1985, 1988; Mann et al., 1988; Kirschvink et al., 1985b) are particularly important; magnetite provides a particularly simple and easy transduction mechanism for virtually any living organism. Magnetite is the only ferromagnetic material known to be a biochemical precipitate, and because its typical interaction energies with the earth's magnetic field are 10⁷ times stronger than any other biological substance, there are many plausible schemes in which it could be used by an organism to provide geomagnetic sensitivity (e.g., Kirschvink and

---

1 From the symposium on Recent Developments in the Study of Animal Migration presented at the annual meeting of the American Society of Zoologists, 27–30 December 1988, at San Francisco, California.
Gould, 1981). In the magnetotactic bacteria, ultrastructural studies reveal the presence of linear chains of magnetite crystals, each held within a specialized membrane-bound subcellular organelle, termed a magnetosome (Balkwill et al., 1980; Gorby et al., 1988). Chains of these structures impart a permanent magnetic moment to the cell and passively rotate it into alignment with the local geomagnetic field (Frankel and Blakemore, 1980). These "biological bar magnet" chains have also been discovered in a magnetotactic protist (eukaryotic algae, Torres de Araujo et al., 1986), and more recently in the ethmoid tissue of salmon (Kirschvink et al., 1985b; Mann et al., 1988). These vertebrate magnetosome chains are remarkably similar in their overall structure and organization to those from the magnetotactic bacteria. Similarly, the discovery that magnetosomes produced by the magnetic bacteria are preserved as fossils in freshwater and marine sediments (Chang and Kirschvink, 1989; Stolz et al., 1986; Petersen et al., 1986) has enabled us to trace the origin of magnetite biomineralization back in the fossil record nearly 2 b.y., which is prior to the origin of the eukaryotic cell (Chang and Kirschvink, 1989). Hence, magnetite biomineralization was probably one of the first matrix-mediated biomineral systems to evolve, and the presence of this ability in a wide range of extant phyla is not unexpected.

Surprisingly, the ultrastructure of these magnetosomes does have a bearing on the question of whether or not higher vertebrates are sensitive to the magnetic field. In terms of their physical properties, they are perfect bar magnets. If the individual crystals were slightly smaller, thermal agitation would cause the magnetic moment to wander relative to the axis of the magnetosome chain, and the organelle would cease to function as a compass. Similarly, slightly larger magnetite crystals would lose their uniform magnetic ordering and break down into multi-domain crystals which would also be ineffective as compass organelles (for a review, see Kirschvink and Gould [1981], Kirschvink [1983], or the Kirschvink et al. [1985a] volume). The alignment of the crystals in a chain also results in the vector summation of their magnetic moments, allowing the production of large magnetic/thermal orientation energies which are necessary for magnetostatic alignment in the geomagnetic field. Without exception, all of the magnetite crystals from magnetotactic bacteria, protists, and fish which have been examined with transmission electron microscopy (TEM) fall within this narrow size range (Kirschvink and Walker, 1985; Kirschvink, 1989). This is not always true of biogenic magnetite produced in chiton teeth, which is unlikely to have a sensory function (Kirschvink and Lowenstam, 1979). The simplest inference from these ultrastructural observations is that the magnetic properties of magnetosomes are maintained by a continual process of natural selection. Hence, the mere presence of magnetosomes in salmon (e.g., Mann et al., 1988) implies that these fish possess at least one important biological process involving geomagnetism. This is not necessarily geomagnetic sensitivity, although the work of Quinn (1980), Quinn et al. (1981) and Quinn and Brannon (1982) clearly supports this hypothesis.

Despite the clear progress during the past 15 years in the understanding of possible mechanisms for geomagnetic sensitivity, it is important to note that there are several fundamental problems which still plague the entire field. First, no behavioral or electrophysiological evidence has yet been presented which uniquely supports any of the many geomagnetic transduction schemes which have been proposed. This dilemma is even true for the elasmobranch fish; although Kalmijn (1978) has shown that elasmobranchs can respond to the magnetic field and possess sensitive electroreceptors, he has not reported the results of experiments to test his proposed mechanism. The role of magnetite in geomagnetic sensitivity is only clear in the magnetotactic microorganisms, where it has been demonstrated by the use of a pulse remagnetization experiment (Kalmijn and Blakemore, 1978; Torres de Araujo et al., 1986). In these experiments, north-seeking organisms can be converted permanently into south-seeking organisms by the
application of a brief and properly configured magnetic impulse. This pulse-reman- 
etization is a unique ferromagnetic effect and cannot be produced in any other fash- 
ion; by itself it is definitive evidence for a ferromagnetic reception system. Thus, the 
development of experimental situations which elicit magnetic north-seeking or south-seeking responses in higher animals would allow us to perform this one definitive test of the ferromagnetic hypothesis of transduction (e.g., Kirschvink et al., 1985b). Later in this paper we will discuss one of our attempts to perform a similar experiment on honey bees in a simple maze.

The second problem concerns the difficulty and reproducibility of many of the behavioral experiments which suggest the presence of a geomagnetic sensory system. A partial list of organisms which supposedly have such a sense includes honey bees (Martin and Lindauer, 1977; Walker and Bitterman, 1985), sharks and rays (Kal- 
mijn, 1978), salmon (Quinn, 1980), tuna (Walker, 1984), eels (Tesch, 1974), hom- 
ing pigeons (Keeton, 1972), migratory birds (Southern, 1978), salamanders and newts (Phillips, 1977, 1986), wood mice (Mather, 1985), and cetaceans (Klinowska, 1985, 1986; Kirschvink et al., 1986). The homing pigeon and honey bee literature is particularly extensive, and the methods used to demonstrate this sensitivity include laboratory conditioning and orientation experiments as well as correlational and exper- imental studies in the field. However, most of the orientation studies depend upon large numbers of observations and elabo- rate statistical analysis of the data, which are highly variable and even now subject to question (e.g., Moore, 1988). Fur- 
thermore, small changes in the apparatus often lead to loss of the orientation effect (e.g., Emlen, 1975). On the other hand, most conditioning studies have obtained negative results. Attempts to replicate at least some of the positive results have ended in failure (Kreithen and Keeton, 1974; Car- man et al., 1987). Finally, field correlation studies such as those on cetaceans and homing pigeons cannot be controlled as well as laboratory-based experiments, and therefore there is difficulty in using these results to guide further work or to test transduction hypotheses.

In view of these difficulties, it is clear that future progress towards answering these questions will depend upon the development of simple, easily-replicated laboratory-based experiments which function well on individuals or small groups of animals. In a series of elegant papers, Walker and Bitterman (1985, 1989a, b, c; Walker et al., 1989) have developed several powerful approaches for conditioning honey bees to magnetic anomalies. In the first of these experiments (Walker and Bitterman, 1985), individual honey bees were trained to feed on sucrose solution set on a shelf before an open laboratory window. During a series of 10 visits, a strong reward (50% sucrose solution) was paired consistently either with the presence or absence of a small vertically-directed magnetic anomaly, and a weak reward (20% sucrose) was paired with the opposite magnetic condition, allowing them to build up a behavioral association between the strong reward and one of the magnetic conditions. Following these visits, the bees were tested for a 10-min period in which they were presented with two targets, both containing tap water, which the animals tasted repeatedly in a vain search for sucrose. Although the bee's tendency to visit the window and look for sucrose was gradually extinguished during this period (hence the term "extinction" test), the choice of targets that it visited yields information concerning stimuli it was able to perceive during the discriminative train- ing. Walker and Bitterman (1985) report that the bees showed a clear preference for the target associated with that magnetic stimulus which had been previously paired with the 50% reward.

In the remainder of this paper, we describe our replication of this basic exper- iment, as well as some slight modifications which make the training paradigm somewhat simpler and produce a better association between the magnetic stimulus and the target. We then describe attempts to alter the paradigm to test for discrimina- tion to a horizontally-directed magnetic anomaly, with little success. Finally, we dis- 


cuss some of our attempts to produce north-
seeking or south-seeking bees in simple maze experiments.

**The Walker-Bitterman Experiment: Geomagnetic Conditioning in Honey Bees**

*Materials and methods*

The subjects were 20 honey bees, *Apis mellifera*, from a single colony which was located on the northwest corner of the North Mudd building of the Caltech campus. In a manner similar to that used by Walker and Bitterman (1985), each experimental animal was selected at random from a group of foragers at a feeding platform equipped with a large jar of 9–12% sucrose solution, perched just outside one of the 3rd-floor, west-facing windows of the Mudd building. Bees were caught at the feeder in a small matchbox and carried to our office, which is also on the top (third) floor, and has three narrow, north-facing windows. We had removed previously the window and frame from the central office window, and replaced it with plastic panels and a 32 cm deep, 44 cm wide, and 20 cm high experimental box as shown in Figure 1A. Bees were released from the matchboxes and placed on a target in the center of the box which had a drop of 50% sucrose, and marked with a spot of colored lacquer. After drinking to repletion, the bees would fly up, hover briefly over the target area, and fly back to the hive. In accord with Walker and Bitterman (1985), the bee usually would return to the window on its own after a few minutes and thereafter continue to shuttle back and forth between its hive and our office window as long as food was available. Sometimes, however, the marked bees would not return to the window and had to be captured again at the feeder and manually placed back on the target. Eventually, most of the bees would return on their own and the experiment could start. Sometimes during the course of an experiment, however, a bee would lead other recruits to the target. As this greatly disturbed the behavioral pattern which we were trying to establish, we were forced to abandon that particular experiment and start again with a fresh animal.

Fluorescent lights were always kept on within the room, to provide a suitable level of illumination within the window box so that we could watch them as well as record their behavior with a video camera. On warm days (peak temperature > 20°C), our honey bees displayed behavior similar to that described by Walker and Bitterman (1985). On cooler days, the time for the bees to return between visits was markedly longer, sometimes only one or two per hour, making completion of the experiment impossible.

Our pretraining and training pattern for both experiments closely followed that used by Walker and Bitterman (1985). During the first few visits, when the bee was carried to the target in the matchbox, the 50% sucrose (or S+ stimulus) was presented under the magnetic field condition (ambient or modified) which would be paired later with the 50% sucrose in the discrimination training. The discriminative training which followed immediately after this consisted of 10 visits, in each of which there was a single target in the middle of the box with the proper sugar concentration. The first discriminative visit was always the S+ condition (20% sucrose), which alternated regularly with the S− condition during the remaining 9 visits and concluded with the S+ prior to the final 10-min extinction test visit. After this final visit, two fresh targets, both containing Pasadena tap water (aversive to these as well as other organisms), were placed over the coil assemblies as shown in Figure 1A for the 10-min interval. Bees, apparently distressed at not finding sucrose, would repeatedly sample the water drop on a target, fly up, land again on one of the two targets, and repeat the process until they gave up. Each direct contact to the water was recorded with a microcomputer programmed to store the position (east or west) and the time to the nearest 0.1 second. For future reference, a videotape record was made for each bee during all of the discriminative training visits and the extinction test. At the end of the 10-min test, the bee was captured, frozen, and the experiment repeated with a new bee from the feeder.

The coil configuration which we used to
produce a small magnetic anomaly differs substantially from those used in the Walker and Bitterman (1985) experiments. Dr. M. E. Bitterman (personal communication) suggested that the behavior might improve if the magnetic stimulus were more tightly localized within the vicinity of each target. Our solution to this suggestion is shown schematically in Figure 1B, and consists of two coplanar, concentric coils of equal dipole moment (area-current), but with antiparallel directions and differing diameter. Close to the center of the coils, the field from the smaller coil dominates and produces the local anomaly. At larger distances, the short-range effects die away and the two dipole moments cancel, effectively confining the anomaly to the 10 cm area surrounding each target. As shown in Figure 1C, we made two of these coil pairs so that one could be located under each of the targets during the final extinction test. Each coil pair was also wound with two separate strands of wire and connected in series as shown on Figure 1C. A DPDT switch was included in the circuit such that it could reverse the current in one of the wire strands and not in the other. Thus, with a single flip of the switch we could move the anomaly from the western to the eastern coil pair (indicated by $L_1$ and $L_2$ on Fig. 1A). The use of the double-wrapped coils in this fashion also forces the currents (and hence the slight amount of heat generated) to be independent of the position of the magnetic anomaly. For all of our experi-
ments we used a standard 1 ampere of current in the circuit generated by a regulated power supply, with the polarity chosen such that the field from the small central coil would add to the ambient field of the window. Walker and Bitterman (1989a, b) have also adopted this coil design in their more recent experiments.

Figure 2A and B shows the pattern of the magnetic anomalies generated by the coils in the two experiments reported below. In the first series, the coils are located under the window box as shown in Figure 1A in the same manner as was used by Walker and Bitterman (1985). The small inner coils produce a downwards-directed dipole moment which adds to the local field. Figure 2A shows the pattern of the total magnetic intensity calculated over a horizontal plane exactly 5 cm above the coils, approximately where the honey bees hover before landing on the target. With 1 ampere of current flowing through the coils, the total intensity of the resultant magnetic field which the bees experience undulates in magnitude with a ‘shape’ (artificially viewed in perspective) which resembles that of a steep pillar surrounded by a shallow moat. The peak anomaly is about 90 µT, or roughly twice as strong as the background level of about 48 µT, and is confined almost entirely to the space over the central target. Figure 2B is a similar calculation for the coils held in an east-west vertical plane with their dipole moment directed horizontally to the north. Note that this configuration yields a pattern similar to the previous one, although the ‘moat’ is slightly more asymmetric. This figure is tilted relative to that of Figure 2A, to emphasize its northwards orientation.

In the first experiment, the targets were made from petri dishes of clear plastic, 5.5 cm in diameter and covered with a thin sheet of green plastic; Walker and Bitterman (1985) used the same type of dish, but sprayed with flat gray paint. As in their experiment, our targets were washed and replaced after each visit from a pool of identical targets in order to randomize extraneous stimuli. In their experiments, Walker and Bitterman (1985) noticed a strong preference in some of their bees for one of the target positions; for this reason they trained the bees alternately in the left vs. the right in an attempt to minimize this interference. In our experiment, however, we placed the training targets in the center of the window box, between the two locations shown in Figure 1, and only used the east and west positions during the 10-min testing interval. We chose to use this somewhat simpler training scheme because our bees displayed less of a position preference than did those in Hawaii.

The second experiment was similar to the first one, except that the coils and targets were mounted on the inside vertical wall. In this case, the targets were made by splitting the green-covered petri dishes into two hemispheres and placing the separate halves on a small plastic ledge which protruded into the window box from the center of the coil pairs.

In both experiments, the presence of the magnetic anomaly could either be paired with the 50% or the 20% sucrose solutions, and during the extinction tests the anomaly could be centered on either the eastern or the western locations. Hence, a minimum of 4 bees are required to balance all possible combinations of anomaly position vs. sucrose concentration. As did Walker and Bitterman (1985), we chose to run at least two bees through each combination, yielding 8 bees per experiment. Plots of the individual responses during the 10-min test intervals for all 16 bees are shown on Figure 3A–P, and summary plots of the data from both experiments, averaged to test for magnetic conditioning, position preference, and anomaly preference are shown in Figure 4. As a final check upon completion of the second experiment, we ran 4 additional bees through the first experiment.

Results

Diagrams in Figure 3 show results from all 16 bees run for these two experiments. Solid lines in each figure show the cumulative number of contacts to the target which had been associated in the discriminative training with the S+ magnetic condition, whereas the dashed lines similarly show the cumulative contacts with the S–
Fig. 2. Pattern of the magnetic anomalies associated with the coil systems used in the conditioning experiments discussed in the text. Calculations of the magnetic field at arbitrary points in space surrounding the coils were made using the method of Stratton (1941) as discussed by Kirschvink (1989), and are summed vectorially with the north, east, and down magnetic components in the window (north = +25 μT, east = +6 μT, and down = +40 μT; 1 μT = 1 microTesla = 1,000 nanoTesla = 1,000 'gamma' = 0.01 Gauss). Calculations were made on a plane 5 cm above the coils, over an area measuring 50 cm by 50 cm and evenly spaced on a grid of 512 x 512 points. Figure 2A is the configuration for the first experiment (similar to that of Walker and Bitterman [1985]) where the coils lie in a horizontal plane underneath the target. Figure 2A is a similar calculation for the second set of experiments discussed in the text, where the coils are mounted on the inside vertical wall of the window box, and have their dipole moment directed due north. The calculations are made along an east-west vertical plane inside the window box, at a distance of 5 cm from the plane of the coils. Direct measurements of the field with a 3-axis fluxgate gave readings within 5% of the calculated values.

target. In virtually every case for the horizontal experiment (Fig. 3A–H) there is a clear preference for the target associated with the magnetic condition (anomaly or ambient) which had been paired previously with the 50% sucrose. In the vertical test, however (Fig. 3I–P), only three of the bees preferred the S+ condition (3J, K, and P), three preferred the S− (3I, L, and M), and the other two (N and O) showed little preference.

In a more formal sense, three things can be tested for in these data: (A) a preference for the targets associated with the magnetic field environments which were paired with the 50% sucrose (S+) during the previous discriminative training, (B) a position preference for either the east or west targets, and (C) a direct preference (or aversion to) the target associated with the magnetic anomaly, independent of whether or not it had been used as the S+ or S− stimulus. The diagrams of Figure 4 show data from both of these experiments averaged according to these groups, after normalizing the response from each bee for the total number of contacts during the 10-min intervals (thereby giving each animal unit weight in the analysis), along with 2-sigma errors around the mean values. Table 1 gives these final averaged results for all 6 of these cases, along with an analysis of variance (ANOVA) to test for significant divergence between the mean values. Only the horizontal test (with the vertically-directed anomaly) yielded a significant dis-
Fig. 3. Extinction test results for the 16 bees used in the horizontal and vertical conditioning experiments. For all bees, the S+ condition is indicated by the solid line, and S− by the dashed line. Bees A–D and I–L were all trained with the S+ reward (50% sucrose) paired with the magnetic anomaly, but during the extinction test the anomaly was on the east for bees A, B, I and J (first column of diagrams) and on the west for bees, C, D, K and L (second column). Similarly, bees E–H and M–P were trained with the S+ reward paired with the ambient magnetic field, and during the extinction test the anomaly was on the east for bees E, F, M and N (third column), and on the west for bees G, H, O and P (fourth column). These diagrams contain all of the information used for calculating the averages given in Table 1, as well as for the 6 summary diagrams of Figure 4.

tinction, and this was very highly significant \((P < 0.001)\) for the conditioned magnetic preference. Our bees did not show preference for either the east or west targets, nor was there any aversion to or preference for the target with the altered magnetic field. Therefore, this is a clear replication of the Walker-Bitterman (1985) experiment.

The failure of our bees in the vertical test to discriminate the northerly-directed anomaly, which in its intensity pattern is similar to that of the horizontal test, was a surprise. Initially we were concerned that, during the two month interval of the experiments, some unknown environmental factor might have changed, interfering with the bee’s ability to sense the magnetic field. However, when we ran a subsequent group of 4 bees through the horizontal test we again obtained clear evidence of magnetic discrimination (F-ANOVA 7.2 with
Fig. 4. Normalized, averaged results from both the horizontal and vertical extinction test experiments, each broken into the 3 groups described in the text and in Table 1. For the S+ / S− group (left pair of diagrams), the solid line shows the average contact frequency of the bees to the target which had the magnetic field (ambient or anomaly) which was paired previously with the S+ condition. The two short dashed lines above and below this line indicate the 2-sigma error boundaries, arbitrarily calculated for plotting purposes at 6 second intervals over the 10-min course of the experiments. Similarly, the S− mean values are plotted with the dash-dot line, with the longer-dashed lines showing the errors. The central pair of diagrams show these results averaged so as to test for an east/west position preference, the solid line (and short-dashed errors) indicating the west preference, and the dash-dot line (and longer-dashed errors) representing the east. Finally, the left pair of diagrams show the same data, averaged to test for a preferential attraction to the magnetic anomaly (solid curve and short dashed errors) or the ambient field (dashed-dot curve with longer dashed errors). Note that the numerical values at the 10-min mark for each of these plots are listed in Table 2, along with results for the ANOVA test of means.

1 and 6 degrees of freedom; P < 0.05). In our experiment, we do not know why bees were not as good at discriminating the horizontally-directed stimulus as they were the vertical one.

**Attempts to Condition Honey Bees in Maze Experiments**

**Background, materials and methods**

Our goal in these experiments was to develop a reward/punishment scheme for training bees to choose a consistent magnetic direction upon exit from a maze, thereby making them exhibit either a north-seeking or south-seeking response. If successful, these then could be tested with the pulse remagnetization experiment and perhaps uniquely test the ferromagnetic hypothesis of magnetic field sensitivity, as was done with the magnetotactic bacteria (Kalmijn and Blakemore, 1978). We preface this section by stating that this work is not yet complete; although bees sometimes show clear preference for magnetic direction in attempts to exit from a simple maze, their response as a function of time is variable, and we do not have adequate control of their behavior. Nevertheless, we shall describe our attack on this problem to date, along with a few of the results which illustrate the problems which remain to be solved. We hope this presentation will stimulate future attempts to solve the behavioral and learning problems, and lead to better attempts with the pulse remagnetization experiment.

Our first design was the T-maze shown
Table 1. Results from the ANOVA test of means for the extinction data of experiments 1 and 2.*

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Group</th>
<th>Mean</th>
<th>Sigma</th>
<th>F-ANOVA</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.</td>
<td>Horizontal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S+</td>
<td>0.667</td>
<td>0.074</td>
<td>81.33***</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>S−</td>
<td>0.333</td>
<td>0.074</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vertical</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S+</td>
<td>0.551</td>
<td>0.129</td>
<td></td>
<td>&gt;0.05</td>
</tr>
<tr>
<td></td>
<td>S−</td>
<td>0.449</td>
<td>0.129</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B.</td>
<td>Horizontal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>West</td>
<td>0.483</td>
<td>0.193</td>
<td>0.118</td>
<td>&gt;0.1</td>
</tr>
<tr>
<td></td>
<td>East</td>
<td>0.517</td>
<td>0.193</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vertical</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>West</td>
<td>0.558</td>
<td>0.126</td>
<td></td>
<td>&gt;0.05</td>
</tr>
<tr>
<td></td>
<td>East</td>
<td>0.442</td>
<td>0.126</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C.</td>
<td>Horizontal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anomaly</td>
<td>0.474</td>
<td>0.192</td>
<td>0.299</td>
<td>&gt;0.1</td>
</tr>
<tr>
<td></td>
<td>Ambient</td>
<td>0.526</td>
<td>0.192</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vertical</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anomaly</td>
<td>0.519</td>
<td>0.139</td>
<td>0.297</td>
<td>&gt;0.1</td>
</tr>
<tr>
<td></td>
<td>Ambient</td>
<td>0.481</td>
<td>0.139</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Sorted to test for (A) preferential sampling of the target associated with the magnetic anomaly which had been paired previously with the 50% sucrose solution (S+ vs. S−), (B) preferential sampling of the east vs. west target on the window box, and (C) preferential sampling of the target associated with the magnetic anomaly. Data analyzed are the final, normalized mean values as plotted on the diagrams of Figure 3. In each case, the F-ratio statistic is standard calculated to test the hypothesis that the mean values for the two groups of 8 are not distinct (implying 1 and 14 degrees of freedom, with large values of F implying rejection). Due to the normalization process, the sum of the group means will always equal 1.00, and the standard deviations (sigma) will always be equal.

schematically in Figure 5A, and was built with the advice and help of Drs. Walker and Bitterman during a brief visit to the Békésy lab at the University of Hawaii in the fall of 1986. The base and 1-cm high walls are made of an opaque black plastic, and have a removable, transparent top. The central floor of the maze contains a narrow, sliding platform which can be moved by pulling a string threaded through a small hole in the front. Bees were trained to drink from a drop of 50% sucrose placed on the center of this platform as described previously, and while they were thus distracted, the platform would be slid gently forward, trapping the bee inside the maze. Upon drinking to repletion, the bee had to choose one of the arms of the T-maze to escape and return to the hive. If she chose the magnetically correct direction, the door would be removed upon her approach (by pulling a string attached to the door), and she would be rewarded with escape. If, however, she chose the wrong direction nothing would happen. The bee would be held within the maze until she retraced her steps and chose the opposite direction, whereupon the door would open allowing her escape. As in the previous experiments, she would usually return from her visit to the hive after a few minutes and land again on the sliding platform, allowing the next trial to begin. Between visits the maze was washed with tap water and dried, so as to remove any visual or odor markings which the bee might have left. During these experiments, the maze was located approximately 20 m due north of the hive and just east of the Békésy lab, and a pair of square Helmholtz coils measuring 1.5 m on a side could be used to invert the horizontal component of the field. The arms of the T-maze were aligned parallel to the north-south magnetic direction, and during half of the trials the field was reversed according to a quasi-random sequence (Gellermann, 1933).

Three of the bees run in this first maze experiment were exposed to a short (1 mS) magnetic impulse with peak intensity of 100 mT (milliTesla), directed antiparallel to a 1 mT bias field, with a portable, battery-powered impulse circuit similar to that described elsewhere (Kirschvink, 1983). A
field of this strength exceeds the coercive force of virtually all known biogenic magnetites. In other experiments with magnetotactic bacteria, this same circuit has been shown to convert north-seeking bacteria into south-seekers, and vice versa (Lin and Kirschvink, unpubl.) as was originally reported by Kalmijn and Blakemore (1978).

Our second maze design is shown schematically in Figure 5B, and our experiments in this operated in a fashion similar to the previous one. Our Pasadena bees were first trained to land on a flat, circular disk which could be raised or lowered within a plastic cylinder like a piston. As a bee drank the 50% sucrose solution, it was lowered gently and an opaque cap placed over the top. The shaft was then connected to a slow motor which spun the bee at 4 rpm in the dark while it finished drinking the sucrose solution. (This was done to minimize the positional preference that bees in the first maze sometimes displayed.) For its escape, the bee would have to choose the correct circular arms of the maze and walk far enough up it to break the beam of an IR emitter/phocell pair. This experiment was also run in the office described above for the extinction tests, with the exits aligned in the east-west direction. We used a system of 12 square coils (4 in each of 3 orthogonal axes) like those described by McElhinny et al. (1971) to cancel the north/south horizontal component of the field, and to apply an equal strength and highly uniform horizontal component along either the east or west direction. A small microcomputer was programmed to monitor the bee’s choice, as well as to control the field direction and record the directional response.

If bees were learning to use either of these mazes properly, we would expect that they would at first make random guesses concerning the exit choice. Upon being punished (delayed from leaving) for making a wrong choice or rewarded (by immediate release) for a correct choice, their exit preference should gradually shift towards the correct direction. We would therefore expect to see some early experimentation, followed by increasingly long strings of choices in the correct direction.

Fig. 5. honey bee maze designs. A. The T-maze used in the experiments of Figure 6A, B and C. Arrows indicate the position of strings used to slide the central platform with the feeding animal into the maze, as well as to remove the exit doors when the bee finally makes a correct choice. B. Cylindrically symmetric maze with rotating floor, used for the bee shown in Figure 6D.

We also need some statistical approach to test for these strings. With the experimental design described above, a bee which is guessing randomly has an equal chance of selecting either magnetic north or south; hence, departures from random guessing can be tested for by use of the simple binomial distribution, in the same fashion as one would use to check flips of a coin. Preference for geographic direction can be tested for in the same manner. As 7 identical choices in a row would be expected once in 128 tries (2^-7 on a one-tailed test), strings of 9 or more (for a 2-tailed test) imply that random guessing is an unlikely explanation. A runs test (e.g., Swed and
Eisenhart, 1943) is slightly less useful in this context, because it is not sensitive to a bias in direction preference and it only checks for grouping based on the total number of runs in the data and not their relative lengths. Hence, a series of short runs followed by an exceptionally long one might not be flagged as significant, even though production of the long run is, by itself, an extremely rare event. We will use both the binomial and runs tests in the following analysis.

Results

Three of the first 8 bees run through the maze of Figure 5A gave hints of a magnetic influence on their response, as shown in Figure 6A, B and C. Directions of first choice for each trial are shown by the tick marks along the bottom of each diagram, plotted with respect to both true north and magnetic north. The cumulative score for the magnetic directions is shown by the graph on each diagram, and the reward direction (that for which the door would open) is indicated along the horizontal axis. After choosing geographic south on 6 of the 8 first visits, the bee of Figure 6A produced a string of 11 out of 12 correct (magnetic north) first choices. The binomial test for these first 20 visits (13 to the magnetic north, 7 south) is not significant, whereas it is for the string of 11 out of 12, with \( P = 0.006 \), two-tailed). A runs test on data from these first 20 visits implies the presence of longer strings than would be expected from chance (\( P = 0.035 \), one-tailed). At this point (between its 20th and 21st visits), the bee was subjected to the 100 mT pulse in the presence of a 0.5 mT antiparallel biasing field and the experiment was continued as before. The next two first choices were both to the magnetic south, followed by an essentially random pattern (neither the binomial nor runs test is significantly different from random, for either the magnetic or geographic directions). Unfortunately, we made a clear mistake in this experimental design—if the ferromagnetic hypothesis is correct, the pulse should have switched the bee’s established preference from magnetic north to south, which was consistent with the animal’s next two choices. However, we forgot to change the reward direction, and the bee was subsequently punished (not released) when it chose the magnetic south direction. In retrospect, it is not surprising that a magnetic preference was lost in the following visits.

Figure 6B shows results from the third bee in this series which, after the first 12 visits, made a string of 11 correct (magnetic north) guesses (for 16 magnetic north choices and 7 south the binomial \( P = 0.092 \); for the 11 in a row it is 0.00098, both on two-tailed tests; a runs test on these first 23 visits shows no significant grouping). Between the 23rd and 24th visits, this bee was also subjected to one of the 100 mT pulses, following which it made magnetic south choices on the next 4 visits, for a total of 7 of over the last 8. We switched the release direction after the magnetic pulse to avoid destruction of the behavior, whereupon the bee continued following a magnetic preference, this time to the south. (As for the previous bee, only the first response after the pulse would contribute formally to a test of the hypothesis, as subsequent responses will be influenced by the reward of having been released on the previous visit.)

Results from the 8th and last bee in this series are shown in Figure 6C, which we were also attempting to train to go to the magnetic north. After 8 of the first 9 choices were to the magnetic south despite continuing punishment (two-tailed binomial \( P = 0.039 \)), we tried another 100 mT pulse in an attempt to flip the bee’s apparent magnetic preference to the north. After the pulse, however, the bee chose geographic north on the next 5 visits, and geographic south on the last 6 (\( P = 0.004 \) on the runs test, one-tailed, implying a non-random arrangement of geographic directional choices). The other 5 bees in the series usually displayed a simple geographic preference, with no suggestion of a magnetic influence on their directional choices.

Subsequent experiments in Pasadena using the maze design of Figure 5B were aimed at trying to enhance the bee’s choice of magnetic direction by reducing or elimin-
Fig. 6. Selected results from individual bees trained in the maze designs shown in Figure 5. Data shown in Figure 6A, B, and C were obtained from the maze of Figure 5A at the Manoa campus of the University of Hawaii; 5 other bees in the same group gave no clear indication of a magnetic exit preference. All three of these animals were exposed to a brief 100 mT magnetic impulse as described in the text. Data from the bee shown in Figure 6D were obtained with the maze design shown in Figure 5B.

inatating the preference for geographic position. Our approach was to deprive the bees of visual cues by having them make their directional choice in the dark, inside a featureless, cylindrical chamber. We slowly rotated their feeding platform during each trial to confuse further the memory of their initial starting orientation. Figure 6D is an example of a magnetically-sensitive bee from this maze. The bee displays a clear preference for magnetic south (37 south choices out of 50, binomial $P = 0.0009$ on a two-tailed test; a runs test is not significant). However, the general behavior in this maze is far from perfect. Long strings of correct responses are often interrupted by periods of apparent random guessing, or in some bees, a switch to equally long strings of incorrect choices. Although these modifications succeeded in reducing the bee’s ability to recognize geographic positions, they did not seem to improve their ability to associate the reward of being released with a magnetic choice. Increasing the punishment for a wrong choice by the application of a small electric shock also did not improve the behavior.

**DISCUSSION**

It is clear that the ability to replicate an experimental result is the acid test of science. This is particularly important in a controversial field, and, because extraordinary claims (such as the presence of geomagnetic sensitivity in animals) require extraordinary proof, positive results should be subjected to independent testing to verify their reality. In this paper, we describe the first independent replication of the honey bee conditioning experiment developed by Walker and Bitterman (1985). With only slight modifications, our results
and conclusions are similar to theirs, and are consistent with many other published reports of geomagnetic sensitivity in honey bees (reviewed extensively by Towne and Gould [1985] and Walker and Bitterman [1985]). With the exception of the horizontal dance alignment first reported by Martin and Lindauer (1977) and replicated by Brines (1978) and Gould et al. (1980), few replications of these other magnetic effects have been attempted. Unlike many of these other experiments, the magnetic extinction test results are not weak. For the horizontal experiment, preference for the target previously associated with the S+ magnetic condition is readily observed in either the raw data of Figure 3(A–H) or the averaged result of Figure 4. The statistical analysis is a window dressing, and it is clear that our results are at least as strong as those reported by Walker and Bitterman (1985).

In addition to confirming the existence of magnetic sensitivity in honey bees, results from our two extinction test experiments imply that this behavior is surprisingly complex. A 90° rotation of the magnetic anomaly and targets from the position shown in Figure 2A to that of 2B was enough to cause the second experiment to fail. We do not know which change was responsible for this difference, but it is clear that, had we started with the second (vertical) experiment, we probably would have abandoned all further efforts. Walker and Bitterman (1989a, b, c) recently have developed a more powerful two-choice paradigm which uses a horizontally-directed anomaly, and provides a level of magnetic discrimination far better than our vertical anomaly results. Hence, our failure in the second experiment was probably the result of a subtle feature in the landing platforms rather than the bee’s inability to detect the stimulus. Consequently, we are not surprised by the failure of numerous geomagnetic conditioning experiments, particularly with animals that have behavior patterns more complex than insects (e.g., Carman et al., 1987 and the entire issue of Animal Behavior 15(2)).

It is intriguing to speculate that the ability of honey bees to respond to the presence of small magnetic anomalies might be telling us something about why honey bees have this geomagnetic sensory system. Although it is clear that very few flowers in nature will be associated as closely with magnetic anomalies as were the targets used in our experiments, magnetic anomalies less than half a meter across can be quite common in some environments. When struck by a lightning bolt, any land surface composed of volcanic or other igneous, crystalline rock will acquire a strong isothermal remanent magnetization (IRM), produced by the magnetic field of the bolt remagnetizing course-grained magnetite in the surrounding material. These anomalies are rather irritating for those of us who also work in the field of paleomagnetism (fossil magnetism in rocks), as they are often powerful enough to deflect compass needles held at wrist height. They are also dense enough in some areas (due to the accumulation of strikes over long periods of time) to serve as a relative dating tool (Kellogg and Shoemaker, 1977). Hence, over large parts of Africa (where honey bees evolved) magnetic anomalies could be useful reference marks for navigation.

The maze experiments are intriguing as well as disappointing. Although only about a quarter of the bees ever exhibit a magnetic exit preference, the magnetic effects when they exist are highly non-random. The data of Figure 6D, for example, might be expected on the average once in perhaps 1,000 experiments, yet we ran less than 20 bees through the apparatus of Figure 5B. Unfortunately, the magnetic results from our second maze design (Fig. 5B) are no better than those from the first (Fig. 5A), implying that none of our manipulations provided better control of the behavior.

Even though we cannot demonstrate that bees will learn magnetic direction in a maze, we can occasionally elicit an unconditioned directional response to the magnetic field. In three cases where the bee displayed magnetic preferences, their behavior patterns changed after exposure to the magnetic pulse (twice the bees switched their magnetic preference on subsequent visits, and one bee lost all magnetic preference
entirely). These results are consistent with the ferromagnetic hypothesis of transduction (e.g., Kirschvink and Gould, 1981); no other transduction mechanism should be influenced permanently by this procedure. This experiment is incomplete, however. Under ideal circumstances, we should use both strong and weak pulses, directed both parallel and antiparallel to the direction of the weaker biasing field. Pulses with a peak intensity less than the coercivity of the magnetite should produce no effect, and the magnetic preference should only change when stronger pulses are directed antiparallel to the bias direction. A reliable conditioning paradigm is essential for this to work, however.

Nevertheless, these maze pulse-re-magnetization experiments add yet another element of support to the ferromagnetic transduction hypothesis for honey bees. This hypothesis is consistent with the presence of biogenic magnetite in the anterior, dorsal part of the abdomen (Gould et al., 1978), the Langevin-shaped accuracy function of the horizontal dance data (Kirschvink, 1981), and by the recent experiment of Walker and Bitterman (1989a, b, c), who found that small magnetized wires mounted on the top front of the abdomen (near the location of the magnetite) interfered with magnetic discrimination, whereas those glued on the thorax or posterior abdomen do not. We also note that the bee of Figure 6D was run in the dark, ruling out the optical pumping suggestion of Leask (1977) unless a biochemical pumping mechanism is also present in the front of the abdomen, far from the visual system. Hence, the magnetite hypothesis is still the simplest transduction model for magnetoreception.

Finally, we repeat that replication is the acid test of science. Even though we were nearly complete novices to behavioral experimentation, we were able to replicate the first Walker-Bitterman experiment. Obviously, much work remains to clarify the transduction mechanism (ferromagnetic or otherwise), yet there is at least one plausible hypothesis which has both behavioral and ultrastructural support as well as clear analogs among a diverse group of microorganisms. We hope that some day the experimental approaches to test for magnetic sensitivity in higher animals are as dependable as those for the magnetotactic bacteria, and can be replicated by advanced high-school students.

ACKNOWLEDGMENTS

We thank Drs. M. M. Walker and M. E. Bitterman for friendly advice on the replication of their conditioning experiment and for help in the design of mazes. We also thank Drs. M. M. Nesson and J. Phillips for critical reviews, and G. Carman for building the coil system used in the Caltech maze experiments. Supported by NSF grants EAR83-51370, EAR86-11512, BNS85-19425 (to MEB @ the University of Hawaii), NIH-BSRG RR07003, and equipment grants from the W. M. Keck and James Irvine Foundations. This is contribution No. 4736 from the Division of Geological and Planetary Sciences of the California Institute of Technology.

REFERENCES

Gould, J. L., J. L. Kirschvink, K. S. Deffeyes, and M.
Honey Bee Magnetoreception


