Chapter 5

The Fifth Force and Eötvös redux

5.1 The rise of the Fifth Force

On January 8, 1986, the *New York Times* ran a front page story with the headline 'Hints of Fifth Force in Nature Challenge Galileo's Findings'. Not only was it on the front page, but there were two graphics under the headline: Galileo dropping two stones of different size from the Leaning Tower of Pisa, and a schematic showing how the interaction of the Fifth Force could make a feather fall faster than a copper coin. Galileo would have been pleased by the publicity—though we suspect not by the rather crudely rendered Tower of Pisa.

The *Times* article was prompted by a modification of Newton's Law of Universal Gravitation, proposed by Ephraim Fischbach, Sam Aronson and their collaborators, that had appeared just two days earlier in *Physical Review Letters* (Fischbach *et al* 1986). In addition to the rapid response by the *Times*, it's noteworthy that the proposed modification was christened as the 'Fifth Force' by John Noble Wilford, the writer of the *Times* story, and not by Fischbach and Aronson[1]. Thus did the *Times* make its lasting imprint on the world of physics. The force was claimed to be composition dependent and thus—contra Galileo—was expected to cause differential rates of fall among different substances.

There had earlier been theoretical speculation about such a modification of Newtonian gravitation at the quantum level, where the change to the gravitational potential would take the form $V = -Gmm'/r\ [1 + \alpha e^{-r/\lambda}]$, which augments the Newtonian potential by the term $\alpha e^{-r/\lambda}$ with α being the strength of the force and λ its range. What caught Fischbach's[2] attention in particular were certain observed

[1] The field theory description of the force involved the exchange of a new, hypothesized particle, the hyperphoton, which differed from the graviton exchanged in the gravitational force. Hence, its description as the Fifth Force. The other four forces are the strong, or nuclear, force, the electromagnetic force, the weak force, responsible for radioactive decay, and the gravitational force.

[2] We will use Fischbach as shorthand for the members of the group, Ephraim Fischbach, Daniel Sudarsky, Aaron Szafer, Carrick Talmadge, and Samuel Aronson.

doi:10.1088/2053-2571/ab3918ch5

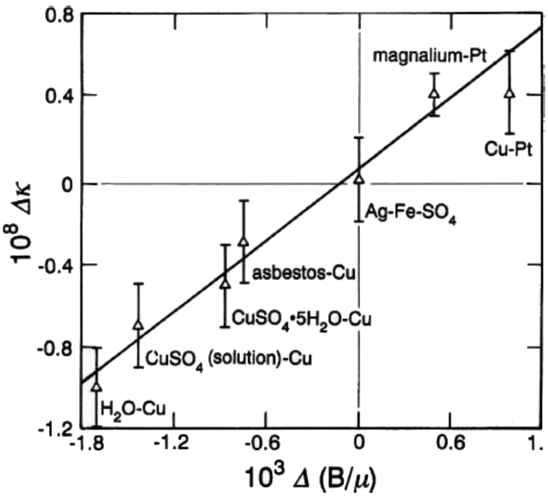

Figure 5.1. Plot of Δk as a function of $\Delta(B/\mu)$. Source: Fischbach *et al* (1986).

CP-violations in K mesons, as well as the difference in measurements of G, the universal gravitational constant, which had been performed in mineshafts and in the laboratory. After further investigation and analysis, Fischbach came to believe that the strength of the gravitational modification arose from a new interaction coupled to the baryon number B of a substance[3] where α, the strength of the Fifth Force, is given by

$$\alpha = -(B_1/\mu_1) \cdot (B_2/\mu_2) \cdot \xi_B$$

where $B_{1,2}$ are the baryon numbers of the interacting objects, and $\mu_{1,2}$ the corresponding masses (in units of atomic hydrogen), and where ξ_B is a defined universal constant.

But how was this to be experimentally tested? Here, Fischbach hit upon the opportunistic and cost effective option of reexamining the data from the original Eötvös experiment with an eye to determining whether there was in fact a discoverable variation in measured gravitational attraction that varied as the product of the ratio of baryon number to mass of the interacting substances. Thus, Fischbach began with a determination of the applicable ratios of baryon number to mass, and then plotted those ratios against the Eötvös data regarding the observed angle of twist with respect to the substance comparisons made (see figure 5.1). Surprisingly, and shocking to many, the comparison as reported in Fischbach *et al* (1986) resulted in a statistical best fit to a linear equation of the slope to be expected on the basis of the Fifth Force hypothesis, as well as the anticipated

[3] For ordinary (baryonic) matter such as considered here, the baryon number is just the sum of the protons and neutrons adjusted to take account of the relative occurrence of isotopes. For the details on how to deal with isotopes and compounds, see Fischbach and Talmadge (1999, pp 19–26).

ordering of the test substances. Furthermore, the range of the parameters needed to fit the data were consistent with the available geological data on G[4].

A more detailed and extensive review of both the motivation for the Fifth Force hypothesis and the reanalysis of the Eötvös data was completed in 1988, where Fischbach concluded:

> Our primary conclusion is that the EPF data taken *by themselves* provide fairly compelling evidence for a new intermediate range force coupling to baryon number or hypercharge. Thus even if it were to turn out that there is no connection between these data and the geophysical and [K meson] analyses which motivated our reexamination of the EPF experiment, the EPF results (if correct) *would be sufficient evidence* for the existence of a 'fifth force'. (Fischbach *et al* 1988, pp 70–1, emphasis added)

Not surprisingly, there was considerable skepticism about the apparently miraculous nature of Fischbach's claim—what one commentator referred to with some gentle irony as a 'gorgeous bibliographical discovery' (De Rujula 1986a, p 761). But once a sign error in Fischback's calculations was identified and rectified, the situation eventually settled and an uneasy consensus developed that Fischback had indeed managed to extract from the Eötvös data what appeared to be confirming evidence for the Fifth Force[5].

But even assuming the soundness of Fischbach's analysis, there remained the possibility of accounting for the data in terms of conventional physics and thus not having to resort to a Fifth Force. Chu and Dicke, for example, proposed such an account in terms of thermal gradients that could exert a force on the test masses. The key fact here, as discussed earlier, was that while the cross sectional area of the test substances was identical for the test masses, the length of the test sample was varied so as to ensure that the mass values were approximately equal. This meant that the surface areas of the different samples used in the Eötvös experiment varied inversely as the density. Thus, there was the possibility that the Eötvös data could be explained in terms of a thermal gradient, and, after determining *post hoc* the values of the relevant parameters by curve fitting, Chu and Dicke were able to generate a coherent and 'reasonably good description of the data' (Chu and Dicke 1986, p 1824).

Fischbach's collaborator, Aronson, made the following objection to the *methodology*—with its *post hoc* determination of parameter values—employed by Chu and Dicke:

[4] For reviews of these constraints, see Franklin and Fischbach (2016, pp 8–9), De Rujula (1986b), and Fischbach and Talmadge (1999, pp 61–3).

[5] For a review of the sign problem and the extensive discussion involved, see Fischbach and Talmadge (1999, pp 8–9) and Franklin and Fischbach (2016, pp 30–1, 180–1). Also, for the imprimatur eventually afforded to Fischbach's reanalysis of the Eotvos data, see De Rujula (1986b, pp 218–20), Fischbach *et al* (1988, p 29), and Franklin and Fischbach (2016, pp 25–6).

> Have we ever understood how, physically, a thermal gradient effect which depends on the length of the samples actually produces a net torque on the balance? … I think *you need a physical hypothesis to test not just a formula with enough parameters to fit the data*. It is true that Dicke doesn't have as many parameters as data, so there is statistical significance to his fit, but why do we ascribe it to thermal effects?[6]

And while it's generally a good idea to have an underlying explanatory theory in hand as opposed to just a 'reasonably good description of the data', it's not always possible to do so. In this regard, it's noteworthy that in the Roll *et al* (1964) solar WEP experiment the test masses were *not* made to look the same externally. This increased the sensitivity of the torsional pendulum to differential disturbing forces of a conventional nature such as temperature variations where 'the most critical effect [of such variations] is that of differential gas pressures associated with *temperature gradients* acting on the balance.' Unfortunately,

> *It seems hopeless to try to understand in detail and quantitatively all of these temperature variation effects*. Fortunately, this is not necessary. It is only necessary to study statistically the correlation between the temperature, its time derivatives, and the balance signal, assuming statistically significant correlations occur. *Once the correlation coefficients are known, the torsion balance signal can be corrected for the temperature effects* by making use of the measured temperature as a function of time. (Roll *et al* 1964, pp 458–9, emphasis added)[7]

There was, however, the following problem with the Chu and Dicke account when applied to the Eötvös data, namely, that the postulated thermal disturbances would have to produce the necessary gradient when the apparatus was rotated 180°, and moreover would have to do so over the entire time period in which the experiment was conducted. Given that requirement,

> It is difficult to imagine how any such source (e.g. a window or radiator) would always produce a gradient with both a fixed direction and a fixed magnitude independent of the time of day or year, over the period of several years during which the experiment took place. (Fischbach 1988, p 48–9)

Still, imagination does not necessarily define the boundaries of the possible[8]. But even assuming a stalemate on this issue, having the hypothesis of the Fifth Force on

[6] As quoted in Franklin and Fischbach (2016, pp 37–8) from an Aronson email, with emphasis added.

[7] The reader at this point may remember that Dicke was a coauthor of this report and a member of the team that developed and ran the solar experiment. He was thus well positioned to offer an account of the Eötvös data in terms of thermal gradients.

[8] For more on Fischbach's objections including the shortcomings of co-opting a property that's monotonically increasing in value to model the only partially monotonic baryon-to-mass ratios, see Fischbach *et al* (1988, pp 47–58).

hand as opposed to just a statistically relevant correlation, led to a consequential and ultimately decisive development: the realization that the *sensitivity* of the Eötvös torsional pendulum could be greatly increased if the experimental apparatus was placed next to a nearby large mass. A simplified analysis will be sufficient to indicate how this sensitivity increase was to be achieved. Since the Fifth Force operates only over a relatively short range, it is reasonable to begin with the assumption of a flat earth, and because the identity of inertia and gravitational mass holds to at least to within 10^{-9}, assume WEP. Under these assumptions, the Fifth Force will act perpendicularly to the Earth's surface and because of that will not exert any torque on a torsional pendulum. In any case, even under more realistic assumptions the displacement of the torsion supporting cord will be very small.

Now introduce a large mass such as a nearby hillside. The Fifth Force originating from this mass will act horizontally, and thus will generate a torque insofar it will affect the test masses as a function of their differing baryon-to-mass ratios. Moreover, since it acts horizontally, its effect will not be diluted by any but the very small angular displacement of the torsional pendulum. Even though highly idealized, this account captures the essence of what's involved in this sort of amplification of sensitivity[9]. If one introduces more realistic assumptions along with reasonable estimates of the density distribution of nearby large masses, it can be shown that the sensitivity could be increased by a factor of 500[10].

Before proceeding further, we wish to emphasize the *conceptual shift* that underlines the above analysis, namely, that an *apparent* violation of WEP (as shown by Fischbach's reanalysis of the Eötvös data) is now construed as the result of a differential, short-range gravitational force that varies as the baryon-to-mass ratios of the test bodies. It is this change in underlying assumptions that creates the possibility—as sketched above—of significant amplification. The question then becomes whether this possibility of amplification will lead to a determinative experimental test of the Fifth Force.

The possibility of such an amplification of sensitivity took on a pivotal role in the debate about the significant of Fischbach's reanalysis of the Eötvös data because on further analysis it became evident that *without the amplifying effect of a large nearby mass* the Eötvös torsional balance would not otherwise be able to detect the Fifth Force. In particular, assuming more *realistic descriptions* of the Earth and the environment of the experimental apparatus indicated that the Eötvös torsional pendulum would not be able to detect the Fifth Force because of subtle distortions due to the centrifugal forces acting on that environment[11].

There is a striking and revealing interaction between Donald Eckhardt and Fischbach on this point. Eckhardt's analysis began with the introduction of a more realistic model of the Earth.

[9] For further elaboration along these lines, see Bizzeti (1986, pp 82–4).
[10] See Bizzeti (1986, p 86), Thieberger (1986, p 2348), Neufeld (1986, p 2345), Milgrom (1986, pp 511–2), and Talmadge *et al* (1986, p 235).
[11] See De Rujula (1986a, pp 760–1, and 1986b, p 217).

> To the zeroth order, the Earth is a sphere held together by gravitation, and the plumb line is directed toward the center of the sphere. To the first order, the Earth is an ellipsoid of revolution held together by gravitation but deformed by the centrifugal forces of its rotation, and the plumb line is not, in general, directed toward the center of the ellipsoid. The ellipsoid is an equipotential surface: *Horizontal gravitational forces … on passive gravitational masses are exactly balanced by opposing centrifugal forces … on inertial masses*. (Eckhardt 1986, p 2868, emphasis added)

Because of this exact balance, it follows that:

> in the absence of local mass inhomogeneities the Eötvös experiment is quite insensitive to any intermediate-range (small compared with the Earth's radius) coupling of any nature. The effect of a local coupling would be to change the magnitudes of the downward forces on the proof masses and on the torsion balance wire, but it would not change the direction of the plumb lines for different composition proof masses; *in effect there would be no horizontal 'Fifth Force' component, so the torsion balance would sense nothing*. (Eckhardt 1986, p 2868, emphasis added)

Fischbach responded by agreeing with Eckhardt that the more realistic assumption 'that the Earth elastically deforms so that its surface lies along an equipotential of [the local acceleration field]' indeed has the consequence that for a short range Fifth Force, 'the Earth as a whole makes no contribution whatever to the EPF anomaly.' (Fischbach *et al* 1986b, p 2869)

If this were the end of the story, it would mean that Fischbach's analysis of the Eötvös data would not be supportive of the Fifth Force hypothesis but would only provide *a measure of the systematic uncertainties involved* in the Eötvös experiment. But all is not lost because as Fischbach further responded:

> [Eckhardt's] observation *has limited practical significance*, since we have already demonstrated … that if the force is of short range then the *dominant contributions* in the EPF experiment will come from local departures from the Earth's geoid. (Fischbach *et al* 1986b, p 2869, emphasis added)

Fischbach's reply succinctly encapsulates the rational response to the sensitivity concerns raised by the use of more realistic models of the Earth, namely, that such higher order niceties can be safely ignored because of the *dominating signal amplification* possibilities afforded by large adjoining masses. Being able to so ignore comparatively small systematic uncertainties exemplifies an important and central feature of experimental practice. That is, of course, where it is possible to do so because of the experimental and underlying theoretical particulars.

Still, Fischbach's response by itself was not sufficient to justify confidence in the Eötvös data as being indicative of the Fifth Force as opposed to being just a measure of the systematic uncertainties. To get the Fifth Force completely off the hook, it would have to be shown that there were *in fact* suitably massive buildings or geological features that would amplify as required the signal sensitivity. After making enquiries, Fischbach received information from Judit Németh regarding the site where the Eötvös experiment was performed. But while suggestive, the information upon analysis was inconclusive. Thus, as conceded by Fischbach, 'neither the magnitude nor the sign of the effective hypercharge coupling can be extracted unambiguously from the EPF data without a more detailed knowledge of the local matter distribution'. (Talmadge *et al* 1986, pp 237–8).

In any case, and putting aside the question of what exactly is shown by the Eötvös data, being able to introduce efficacious horizontal components of gravitational force meant that there was a *specific* and *open possibility* for *significant improvement* in the reliability and precision of the results of *replications* of the Eötvös experiment. Where, of course, these possibilities were contingent on the truth of the Fifth Force hypothesis. This put the onus on the experimentalists to take advantage of the amplification possibilities of nearby mass inhomogeneities and devise experiments that were correspondingly more sensitive and decisive. Consequently, and to return to the themes of our Introduction, there emerged a clear career and publication bias, this time, *in favor* of replication. Moreover, this bias was backed up by an accumulated and substantial combination of experimental expertise and theoretical underpinning—a combination that was capable of substantiating the *credibility* of claimed replications.

5.2 Its fall

In fact, it didn't take long for the first replications to be reported at the Moriond Workshop (Thieberger 1987a, Raab *et al* 1987), and then subsequently published as adjoining papers in *Physical Review Letters* (Stubbs *et al* 1987, Thieberger 1987b). The experiments were conducted using rather different apparatus and methods. This diversity meant that an agreement of result would be strongly confirming or disconfirming of the Fifth Force hypothesis because of the unlikely possibility that the different systematic uncertainties involved would all conspire to yield the same result.

There was, however, disappointment on this score. The results were not in agreement. Peter Thieberger's results supported the existence of a Fifth Force, whereas the results of the whimsically named Eöt-Wash group[12] found no evidence for such a force. The two experiments *considered together* thus constituted a failed replication. The problem then was to determine what was to blame for this difference of result, the systematic uncertainties or the non-existence of the Fifth Force.

Thieberger's experiment was especially noteworthy because it avoided the complication of having to take account of both gravitational and centrifugal forces since it relied only on the differential Fifth Force effect due to the Palisades cliff in

[12] The name is a play on the pronunciation of Eötvös (Uht-vush) because group was located at the University of Washington.

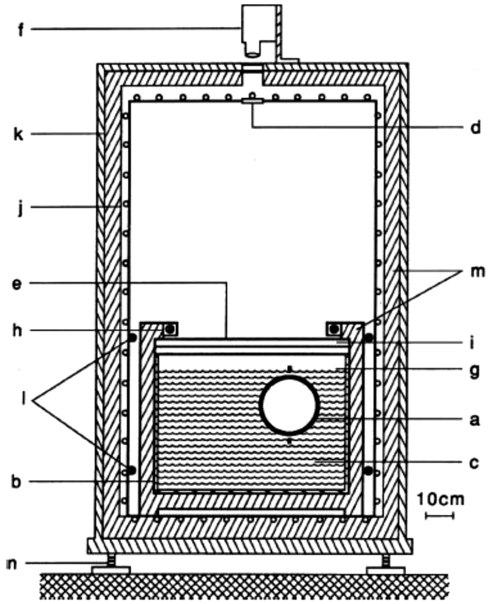

Figure 5.2. Schematic diagram of the differential accelerometer used in Thieberger's experiment. A precisely balanced hollow copper sphere (a) floats in a copper-lined tank (b) filled with distilled water (c). The sphere can be viewed through windows (d) and (e) by means of a television camera (f). The multiple-pane window (e) is provided with a transparent x–y coordinate grid for position determination on top with a fine copper mesh (g) on the bottom. The sphere is illuminated for 1 s per hour by four lamps (h) provided with infrared filters (i). Constant temperature is maintained by means of a thermostatically controlled copper shield (j) surrounded by a wooden box lined with styrofoam insulation (m). The mumetal shield (k) reduces possible effects due to magnetic field gradients and four circular coils (l) are used for positioning the sphere through forces due to ac-produced eddy currents, and for dc tests. Source: Thieberger (1987a).

New Jersey. This simplification was possible because Thieberger employed a differential accelerometer in which a copper sphere was submerged in water such that the center of mass of the copper sphere corresponded to that of the displaced water. The experimental apparatus is shown in figure 5.2.

As can be seen, and as reflected in Thieberger's analysis, the efficacious differences in the forces acting on the copper and the water would be exclusively those originating from the Palisades. Assuming, of course, that the systemic uncertainties were ignored at least for the time being. Thus, given the Fifth Force hypothesis, the copper sphere was expected to accelerate through the water where the Fifth Force contribution could be determined through an application of the Stokes equations.

Thieberger's results, taken over five days, are shown in figure 5.3. He found a velocity of (4.7 ± 0.2) mm h^{-1} in the y-direction and a velocity of (0.6 ± 0.2) mm h^{-1} in the x-direction, where the y-axis points approximately east, away from the cliff. The measured y-velocity corresponded to an acceleration of $(8.5 \pm 1.3) \times 10^{-8}$ cm s^{-2} or a force of $(4.2 \pm 0.6) \times 10^{-4}$ dynes. These values were used to set limits on the values of $\alpha\lambda$ for the parameters of the Fifth Force and it was determined that for

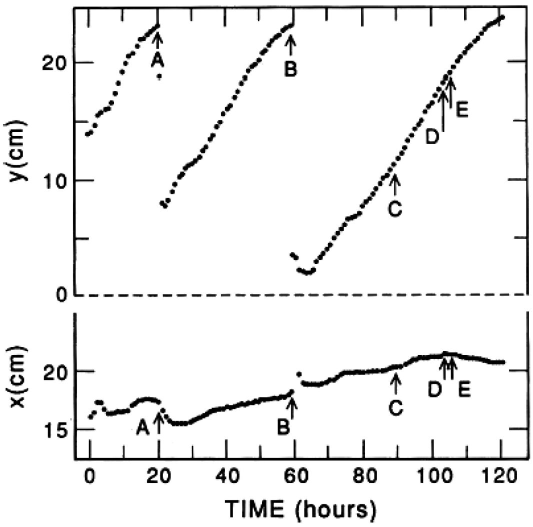

Figure 5.3. Position of the center of the sphere as a function of time. The *y* axis points away from the cliff. The position of the sphere was reset at points A and B by engaging the coils shown in figure 5.2. Source: Thieberger (1987a).

5 m < λ < 100 m, $\alpha\lambda \approx (1.2 \pm 0.4)$ m, which was consistent with Fischbach's estimate based on the Eötvös data.

In addition to having a simplified ensemble of forces at play, Thieberger's apparatus also provided a way of side stepping the following problem, which we noted earlier, with the original Eötvös experiment, namely, its 'lack of a suitable control':

> There is no way of turning off the centrifugal force field of the Earth. Hence, there is no over-all zero check upon the performance of the torsion balance. (Roll *et al* 1964, p 446)

But with Thieberger's apparatus, a 'suitable control' was possible and an 'over-all zero check upon the performance' could be made. This is because the horizontal gravitational forces due to the Palisades cliff (the principal source of the differential forces involved) could be in effect turned off by simply moving the apparatus to a different, cliff-free location. So, Thieberger moved his differential accelerometer so that measurements were made 'in the absence of a cliff but under otherwise similar conditions.' (Thieberger 1987a, p 1068).

While not explicitly stated by Thieberger, the methodological principle here is that if systematic errors were responsible for the non-null result recorded in the presence of the Palisades, then moving the apparatus away from the cliff should result in velocity components that were essentially unchanged. The result, as hoped for, was that significantly smaller velocity components were recorded. Thieberger, however, and surprisingly so, did not expressly note that these values provide a *reasonable estimate* of the systematic uncertainty. That such a reasonable estimate was so provided was explicitly noted by Paul Keyser though he suggested as well that the resulting estimate was rather large:

> … in a symmetric environment (i.e. far from the cliff) where any anomalous force should be zero, a force equal to ~ 1/3 of the anomalous force was still observed… This suggests a large part of the anomalous force may be systematic. (Keyser 1989, p 2332)

> Thieberger replied that:

> Not only were these velocities *much smaller* than during the main measurement, but they also appeared to be *randomly directed*. (Thieberger 1989, p 2333, emphasis added)

In other words, since the systematic uncertainties were 'much smaller' as well as 'randomly directed', they were not sufficient to disturb the essentially positive result of the experiment.

In addition to 'turning off' the horizontal gravitational force components, Thieberger employed what is now a standard technique for dealing with certain systematic known as 'background effects.' Here, the basic methodology is to impose a much stronger background source signal than was initially present in the experimental apparatus. One then shows that the effect of such amplification is of little or no consequence with respect to what is intended to be measured. This, as the reader may recall, is *exactly the methodology employed by Newton* when he amplified the effect of the postulated Cartesian aether and demonstrated that the systematic error caused by such an aether would be of no consequence when considering his pendulum experiments. So, for example, Thieberger elevated temperature differences, displaced leveling by 'over ten times,' and introduced a fourfold increase in illumination. All these amplifications of background signal were 'without appreciable effect.'

In a variant on this methodology, and to guard against possible instrumental asymmetries, data was taken for 28 h after the apparatus had been rotated 90° to maximize the asymmetry. The measured y-velocity was (4.5 ± 0.5) mm h^{-1}, consistent with the initial result. In other words, the apparatus was placed in an orientation that would in a sense be amplifying because the asymmetries, if present, would have a correspondingly different effect.

Thieberger also claimed that untoward effects due to 'residual dipole moment and higher multipole moments, electrostatic and magnetic forces, electromagnetic radiation, surface tension and its temperature dependence, convection currents, vibrations, temperature gradients, and Brownian motion' could be similarly 'largely ruled out' as would soon be 'described elsewhere'. (Thieberger 1987b, p 1068) While Thieberger never got around to describing 'elsewhere' how these additional systematic uncertainties were to be dealt with, he did add a late publication note that took into account 'that the Coriolis force on the sphere was not totally negligible'. Taking that force into account, however, resulted in a 'slightly less satisfactory agreement',

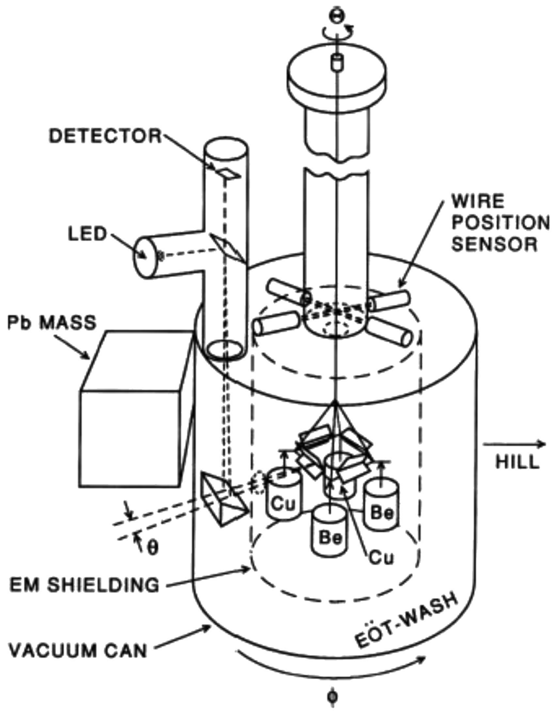

Figure 5.4. Schematic view of the University of Washington torsion pendulum experiment. The Helmholtz coils are not shown. From Stubbs *et al* (1987).

which however was rectified by using a more appropriate measure of the effective cliff orientation (Thieberger 1987b, 1068–9)[13].

The Eöt-Wash paper was published next to Thieberger's paper in *Physical Review Letters* and presented a conflicting result. The experimental group remarked that they had been motivated to perform the experiment by the anomalies in the geophysical measurements of G and by Fischbach's reanalysis of the Eötvös experiment which had shown a composition dependence. More informally, Eric Adelberger, a group leader, remarked that the idea of testing a fundamental law with a comparatively inexpensive and conceptually simple experiment was an 'intriguing possibility' (Adelberger, private communication).

The experimental apparatus is shown in figure 5.4 and consisted of a freely oscillating torsion pendulum containing two beryllium and two copper test bodies arranged as shown. Beryllium and copper had B/μ ratios of 0.99865 and 1.00112, respectively, and thus maximized the difference in B/μ ratios. The four-body pendulum was located on the side of a hill on the University of Washington campus, which provided the local mass asymmetry needed for an observable Fifth Force effect. Therefore, it was expected that:

[13] For more details, see Thieberger (1987a, pp 581–4).

> The balance will experience a torque if the Be and Cu test bodies are attracted differently to nearby matter (the 'sideward' pull of the hill) than to distant matter (the 'downward' pull of the Earth). (Stubbs *et al* 1987, 1070)

As it stands, the comment is somewhat ambiguous but obviously is to be understood in terms of the analysis (reviewed above) of the increase in sensitivity that would result—assuming the existence of the Fifth Force—because of the proximity of a large mass such as the nearby hill.

Echoing Newton's preparation of his pendulums so that they were 'exactly like each other with respect to their weight, shape, and air resistance' (Newton 1999, p 807), the Eöt-Wash group following suit stated that:

> We minimize false signals by designing the test bodies to *appear identical in all respects except for baryon content*. Each body was a cylinder 1. 908 cm high and 1.905 cm in diameter and had a mass of 10.04 g. The external dimensions of the bodies were identical to within ± 0.0025 cm and their masses were equal to ± 4.6 mg. The difference in density between Be and Cu was accommodated by fabrication of the Cu bodies as cylindrical shells fitted with endcaps. Care was taken to assure that the centers of mass of the hollow bodies coincided with their geometrical centers. (Stubbs 1987 #758, p 1071, emphasis added)

In addition, the apparatus was mounted on a turntable and slowly rotated at a constant speed. This minimized the stress on the torsion cord. The experimental procedure was to take deflection angle data at 90° increments for at least 10 complete revolutions of the apparatus. The sinusoidal variations in deflection angle to be expected as the apparatus was rotated were calculated on the basis of an estimate of the Fifth Force contribution from the nearby hill. The superposition of experimental results and sinusoidal expectation are reproduced in figure 5.5.

In short, the result of all the careful preparation was that there was no change in torsion as predicted by the Fifth Force hypothesis:

> *There is no apparent signal from external sources*, but we do observe an offset of ≈ 4 μrad. This systematic effect is presumably due to an imperfection in the can rotation drive and to thermal gradients fixed in the laboratory frame… Our results rule out a unified explanation of the apparent geophysical and Eötvös anomalies in terms of a new baryonic interaction with $10 < \lambda < 1400$ m and make it highly implausible that the systematic effects in the Eötvös data are due to a new fundamental interaction coupling to B. (Stubbs 1987, pp 1071–2, emphasis added)

The theoretical predictions for the Fifth Force shown in figure 5.5 were based on a value of $\alpha = 0.001$ with $\lambda = 100$ m. At the time, the best estimate of α was ≈ 0.01, so that the predicted effect shown was, in fact, underestimated by a factor of ten. Eric Adelberger stated that the group wanted to be as conservative as possible in their evaluation of the Fifth Force hypothesis (private communication). A *more realistic*

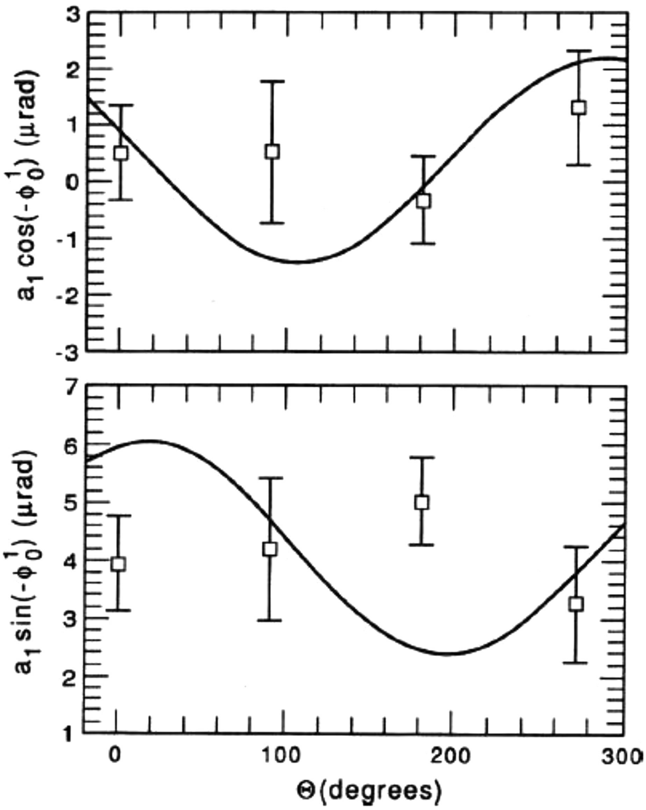

Figure 5.5. Detection signal as a function of θ, the variable angle of the stage. The theoretical curves correspond to the signal expected for $\alpha = 0.001$ and $\lambda = 100$ m. Source: Stubbs *et al* (1987).

estimate of the predicted effect was presented at the Moriond Workshop (Raab *et al* 1987) and is shown in figure 5.6 where the assumed value of α for this graph was 0.01. As can be readily seen, the null result is now clearly evident. The experimenters set constraints on α using local topography and the proposed Fifth Force. At the 1 σ level, these were $|\alpha| < 2 \times 10^{-4}$ for $250 < \lambda < 1400$ m and $|\alpha| < 1 \times 10^{-3}$ for $10 < \lambda < 250$ m.

As with Thieberger's paper, a considerable portion of the Eöt-Wash paper was devoted to describing the apparatus, its operation, and its calibration. Thus, in addition to 'designing the test bodies to appear identical in all respects except for baryon content,' the Eöt-Wash group reported that it:

> paid particular attention to systematic effects that could either produce a false signal or possibly cancel a true signal. The most important sources of such errors are (1) departures from four-fold symmetry in the torsion pendulum…; (2) deviation of the can axis of rotation from true vertical [i.e. tilt]; and (3) thermal gradients across the apparatus. (Stubbs 1987, p 1071)

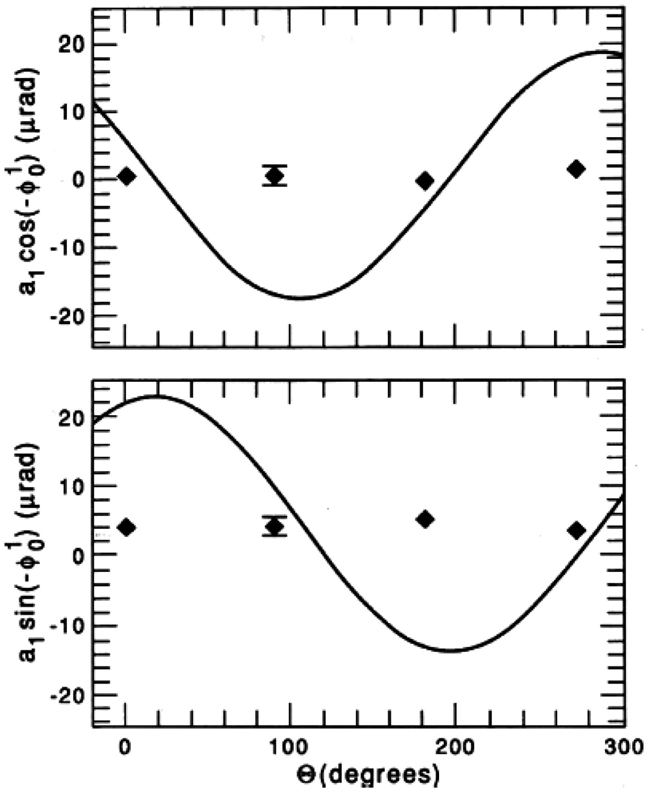

Figure 5.6. Detection signal as a function of θ. The theoretical curves correspond to the signal expected for $\alpha = 0.01$ and $\lambda = 100$ m. From Raab *et al* (1987).

None were large enough to affect their results. Since there was a null result, there was no reason to shut off the horizontal gravitational force by moving—as Thieberger had done—the apparatus away from the hill. Unless, that is, the systematic uncertainties were suspected—and they were not—to have exactly compensated for and rendered invisible the Fifth Force effect.

In sum, there were disparate results: with one experiment indicating a positive effect, and the other a null result. Moreover, there were no readily available tiebreakers since both Thieberger and the Eöt-Wash group gave well-considered arguments for the credibility of their results. In particular, they checked for many of the same confounding background effects including those due to magnetic fields, instrumental asymmetries, thermal gradients, and leveling errors. Yet the results disagreed. Because the experiments used very different types of apparatus, a floating sphere for Thieberger, and a torsion pendulum for the Eöt-Wash group, there might

Table 5.1. Data taken from Adelberger *et al* (1991).

$\lambda = 1000$ m	$\lambda = \infty$	Reference
$(1.1 \pm 4.3)10^{-6}$	$(1.3 \pm 5.1)10^{-10}$	Niebauer *et al* (1987)
$(0.4 \pm 1.1)10^{-5}$	$(0.4 \pm 1.3)10^{-9}$	Kuroda *et al* (1990)
$(1.1 \pm 6.6)10^{-6}$	$(1.3 \pm 8.0)10^{-10}$	Kuroda *et al* (1990)
$(-0.2 \pm 1.2)10^{-5}$	$(-0.2 \pm 1.4)10^{-9}$	Kuroda *et al* (1990)
	$(-1.3 \pm 1.5)10^{-11}$	Roll *et al* (1964)
	$(3.0 \pm 4.5)10^{-13}$	Braginskii *et al* (1972)
$(1.2 \pm 0.3)10^{-7}$	–[d]	Boynton *et al* (1987)
$(1.4 \pm 4.2)10^{-9}$	$(-0.2 \pm 1.0)10^{-11}$	Heckel *et al* (1989)
$(-5.1 \pm 5.1)10^{-9}$	$(-0.5 \pm 1.3)10^{-11}$	Heckel *et al* (1989)
$(-1.1 \pm 2.0)10^{-7}$	$(1.8 \pm 12.9)10^{-9}$	Fitch *et al* (1988)
$(-7.2 \pm 7.6)10^{-4}$	$(-7.2 \pm 7.6)10^{-4}$	Speake *et al* (1988)
$(-3.2 \pm 3.7)10^{-4}$	$(-3.2 \pm 3.7)10^{-4}$	Speake *et al* (1988)
$(-0.7 \pm 1.4)10^{-5}$	$(-0.7 \pm 1.4)10^{-5}$	Bennett (1989a)
$(-0.4 \pm 3.9)10^{-6}$	$(-0.4 \pm 3.9)10^{-6}$	Stubbs *et al* (1989b)
$(-1.4 \pm 0.9)10^{-6}$	$(-1.4 \pm 0.9)10^{-6}$	Cowsik *et al* (1990)
$(-1.1 \pm 1.2)10^{-6}$	$(-1.1 \pm 1.2)10^{-6}$	Nelson *et al* (1990)
$(-5.1 \pm 1.7)10^{-6}$	–[d]	Thieberger (1987b)
$(0.0 \pm 1.1)10^{-7}$	$(0.0 \pm 1.4)10^{-9}$	Bizzeti *et al* (1989b)

have been some crucial difference between the two types of apparatus that accounted for the conflicting results. But no clear candidate for such a difference emerged[14]. Indeed, after several years of such scrutiny, and even to this day, no one has found a tie-breaking error in either experiment or its analysis.

With the situation thus deadlocked, there was ample motivation for yet more attempted replications, and there were many, including a combined half-ring torsional balance, updates of the Tower of Pisa experiment, and another floating sphere experiment. In short, the result was that none of the replications agreed with Thieberger's result[15] (see table 5.1). In this regard the experiments of Bizzeti and his collaborators were particularly influential because they (Bizzeti *et al* 1988, Bizzeti *et al* 1989b) used the same type of experimental apparatus as Thieberger had, and revealed no Fifth Force effect. Especially compelling because if Bizzeti's experiments had indicated a positive result, there would have been coordination of result with the *type* of apparatus used. The problem then would have been to determine whether, for example, the floating-ball or torsional pendulum apparatus had the upper hand when it came to the treatment of systematic uncertainties. But that problem was avoided once Bizzeti's results were combined with the other null results. In addition,

[14] See Fischbach and Talmadge (1999, pp 172–5, 146–55) for a more extensive review of the Thieberger and the Eöt-Wash experiments.

[15] These experiments searched for a composition-dependent Fifth Force and, in some cases, a possible distance dependence. For details, see Franklin and Fischbach (2016) and Fischbach and Talmadge (1999, pp 146–77).

Bizzeti conducted a *series* of experiments with corresponding improvements made in response to the experimental difficulties *specific* to the use of a differential accelerometer[16].

In light of the accumulation of null results, and in response to an objection from Paul Keyser regarding the likely existence of a certain form of causally efficacious convection current, Thieberger issued a qualified concession:

> The observed motion *could indeed have been due* to ordinary forces. Unanticipated spurious effects can easily appear when a new method is used for the first time to detect a weak signal. Neither the title nor the text of [my 1987 report] contains a claim to the discovery of a new force … After the initial suggestion to perform this type of measurement close to cliffs and after the first such experiment, several other experiments now have been reported, most of them conducted with more conventional instruments. Even though the sites and the substances vary, effects of the magnitude expected have not been observed. Therefore … it now *seems likely* that some other spurious effect *may have caused the motion* observed at the Palisades cliff. (Thieberger 1989, p 2333) (emphasis added)

At the 1990 Moriond workshop, attended by many of the researchers working on the Fifth Force, Orrin Fackler summed up the situation when he stated quite emphatically and without qualification: 'The Fifth Force is dead.' No one present disagreed[17].

While Fackler's death pronouncement and the agreement of those at the workshop were warranted, we wonder what exactly is the justification for this sophisticated variant of crowd sourcing? While clearly rational under the circumstances, the excommunication in effect of the Thieberger experiment is nevertheless unsettling because, as already noted, there was not a convincing explanation of why things had gone wrong. The excommunication was thus based on, to use a concept from the criminal law, purely circumstantial evidence precisely because there was not a direct and agreed diagnosis of the experimental failure. Similarly, for Fischbach's analysis of the Eötvös data. If not explained by the now deceased Fifth Force, then by what? So too here, no convincing accounts have appeared[18].

One possibility for explaining the effect that Fischbach and his collaborators found in the Eötvös results concerns the choice of data and data analysis procedures.

---

[16] (Bizzeti 1987, Bizzeti *et al* 1988, Bizzeti *et al* 1989a, Bizzeti *et al* 1989b, Bizzeti *et al* 1990).

[17] Franklin was present at the conference.

[18] See Fischbach and Talmadge (1999, pp 213–4), Franklin and Fischback (2016, pp 204–8), and Hall *et al* (1991) for the curious and unexplained correlation of the baryon-to-mass ratios and charge-to-mass ratios. In a recent communication, Fischbach informed us that his current appraisal is that the 'Hall's 1991 paper only accounts for the EPF (Eötvös) correlation in a crude way and so is not likely to be correct. Moreover, it is not highly motivated theoretically.' More generally, 'it would be very difficult for any classical variable' to achieve the 'level of precision' required in order to duplicate what had been achieved in the original reports by Fischbach *et al*. Thus, Fischbach's 'working hypothesis at present is that the effect really is proportional to B/μ, but that this contribution is amplified, or catalyzed, by some other physics which we are not yet taking into account.'

Eötvös conducted his experimental determination on essentially eight substances, four of which were compared with the platinum standard, while four were compared with a copper standard. In his final analysis, the four comparisons with copper were 'converted' into comparisons with platinum by means of his water to copper, and copper to platinum experimental values. This conversion, however, had the effect of reducing the significance of the now converted data. Thus, as noted by Fischbach: 'The effect of this combining say $\Delta k(H_2O\text{-}Cu)$ and Δk(Cu-Pt) to infer $\Delta k(H_2O\text{-}Pt)$ is to reduce the magnitude of the observed nonzero effect [for water and platinum] from 5σ to 2σ.' (Fischbach *et al* 1986, p 6). In particular, $\Delta k(H_2O\text{-}Cu) = (-10 \pm 2) \times 10^{-9}$ and $\Delta k(\text{Cu-Pt}) = (+4 \pm 2) \times 10^{-9}$, respectively. Adding them to obtain $\Delta k(H_2O\text{-}Pt)$ gives $(6 \pm 3) \times 10^{-9}$.

In order to avoid this reduction in significance Fischbach used the original, unconverted data. Figure 5.7 shows both the summary reported by Eötvös as well as Fischbach's reanalysis, along with best-fit straight lines for both sets of data separately (this is Franklin's analysis). Although several of the experimental uncertainties have increased, due to the calculation process, the lines have similar slopes. The major difference is in the uncertainty of the slopes. If one looks at the 95% confidence level, as shown separately for the Fischbach and Eötvös data, respectively, in figures 5.8 and 5.9, one finds that at this level the published, tabulated Eötvös data are, in fact, consistent with no effect, or a horizontal straight line. This is not true for the Fischbach reanalysis. The startling and unexpected result found by Fischbach and his collaborators may be an artifact of the data analysis. We are not in any way suggesting that Fischbach *et al* selected their procedure to get their positive result.

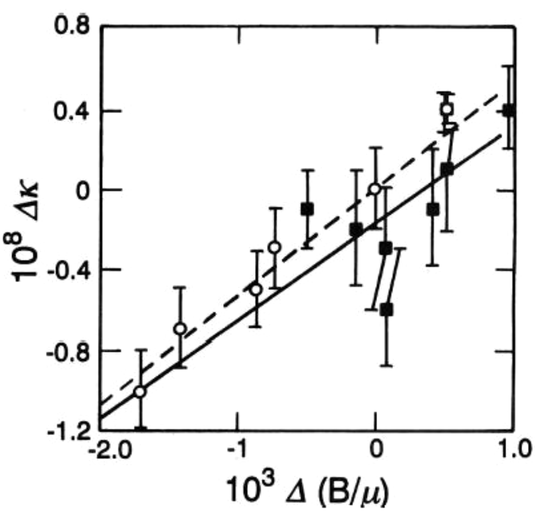

Figure 5.7. Plot of Δk as a function of $\Delta(B/\mu)$. Circles are data from Fischbach *et al* (1986a) Squares are the final summary from (Eotvos *et al* 1922). The dashed line is the best fit straight line to Fischbach's data. The solid line is the fit to the Eötvös data.

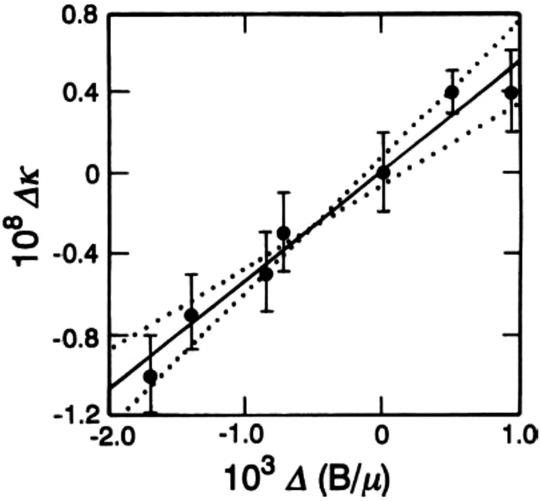

Figure 5.8. Plot of Δk as a function of $\Delta(B/\mu)$ from Fischbach *et al* (1986a). The best fit line along with the 95% confidence level fits are shown.

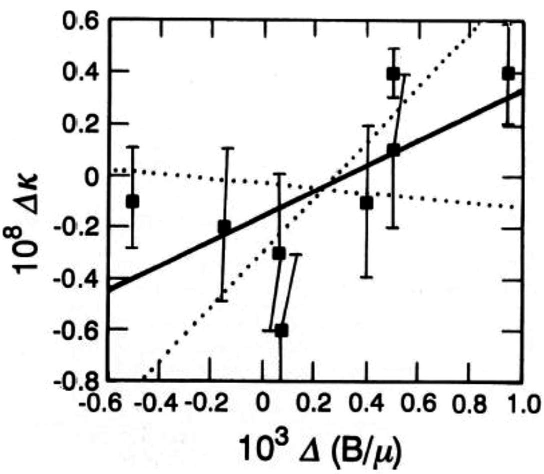

Figure 5.9. Plot of Δk as a function of $\Delta(B/\mu)$. The data are from Eötvös (1922). The best fit straight line along with the 95% confidence level fits are shown.

While any such concern about the circumstantial nature of the evidence may be dismissed as overwrought, it does serve the purpose of highlighting that there are limitations—both practical and theoretical—to what can be fully explained. In such cases, crowd sourcing such as here exemplified will have to do. Of course, while in some respects circumstantial, that crowd sourcing was sophisticated and scientifically informed. Moreover, it is precisely the *converging agreement* that occurred along with the *increasingly sophisticated standards of the appraisal* of theory and experiment that justified the diagnosis that the Fifth Force was dead.

Still, there was apparently some residual unease with the pronouncement of the death of the Fifth Force, and some experimentation continued afterwards though with much of the motivation coming from the perceived need to further nail down the accuracy of WEP and, as well, the possibility of discovering (consistent with increasingly more accurate determinations of WEP) some variant form of the Fifth Force.

5.3 Tests of the weak equivalence principle

Despite what one might have expected, the demise of the Fifth Force did not cause any lack of enthusiasm among experimentalists for developing ever more accurate and expansive tests of Newtonian gravitation. If anything, there was an upsurge in interest. There were essentially three reasons for this.

First, even though the Fifth Force was no longer on the scene, the underlying theoretical basis for suspecting that there might be modifications to Newtonian gravitation was still largely intact. On the experimental side there was the accumulated expertise developed as a result of the extensive testing of the Fifth Force hypothesis where this expertise was at the ready for continuing investigation of such modifications. Thus, even though refuted, the Fifth Force nevertheless served to generate interest in further research regarding non-Newtonian gravitation. Fischbach took both solace and satisfaction in this positive result:

> My guess is that searches for deviations from Newtonian gravity would have had a much more difficult time becoming part of mainstream physics, had it not been for the Rencontres de Moriond and the credibility they lent to such efforts. In addition to the meetings themselves, and the opportunities they provided for interactions among the participants, the Proceedings from each meeting played an important role by collecting together many of the early experimental results and theoretical ideas. (Franklin and Fischbach 2016, p 194)

Second, there was the continuing fundamental importance of WEP for the General Theory of Relativity, and, in particular, for attempts to integrate it with the standard model of particle physics:

> The equivalence of gravitational mass and inertial mass is assumed as one of the most fundamental principles in nature. Practically every theoretical attempt to connect general relativity to the standard model allows for a violation of the equivalence principle. Equivalence-principle tests are therefore important tests of unification scale physics far beyond the reach of traditional particle physics experiments. The puzzling discoveries of dark matter and dark energy provide strong motivation to extend tests of the equivalence principle to the highest precision possible. (Schlamminger *et al* 2008)

Third, it was realized that experimental examinations of WEP could be understood and used both as *tests* of WEP and as *methods of discovery* for new forms of non-Newtonian gravitation:

> The universality of free fall (UFF) asserts that a point test body, shielded from all known interactions except gravity, has an acceleration that depends only on its location. The UFF is closely related to the gravitational equivalence principle, which requires an exact equality between gravitational mass m_g and inertial mass m_i and therefore the universality of gravitational acceleration. Experimental tests of the UFF have *two aspects*—they can be viewed as *tests* of the equivalence principle or *as probes for new interactions* that violate the UFF. (Su *et al* 1994, p 3614, emphasis added)

Our aim in this chapter is to explain what makes these two aspects of experimental testing possible. That is, how experimental tests of WEP can be used not only to confirm or disconfirm but also to *probe and search* for new forms of non-Newtonian gravity. In order to understand what makes this dual usage possible we'll need to return to what is known as the Yukawa potential which was briefly discussed above in chapter 5 when we dealt with the motivation and basis for Fischbach's development of the Fifth Force hypothesis.

The Yukawa potential is a theoretically based formalism of convenience and choice used to represent the principal elements involved in a modification or supplementation of Newtonian gravitation[19]:

$$V(r) = -G\,\frac{m_T m_S}{r}\left[1 + \alpha\left(\frac{q_T}{\mu_T}\right)\left(\frac{q_S}{\mu_S}\right)e^{-r/\lambda}\right].$$

Here, the Newtonian gravitational potential is augmented where α is the strength of the augmentation relative to that of gravity, λ the range of the augmentation, and r the distance between the test mass and the gravitational source also known as the attractor. The term q/μ refers to the 'charge' per atomic unit of mass where what is taken to be the 'charge' is the property suspected as being responsible for the augmentation. The leading candidates here are the baryon number B, the lepton number L, and B—L as well as other combinations of B and L. The subscripts T and S refer, respectively, to the test mass and the source of the gravitational attraction.

Experiments designed to determine the difference in acceleration of masses of different substance caused by a known attractor involve, in their analysis, the comparison of the Yukawa potentials for the different test masses. In short, any difference in acceleration will be reflected in a difference of Yukawa potentials. Expressed formally, this difference is:

[19] This application was originally suggested by Fujii (1971). For more background, see Fischbach and Talmadge (1999, pp 3, 12–4).

$$V\Delta(r) = -G\,\frac{m_T m_S}{r}\left[\alpha\left(\Delta\frac{q}{\mu}\right)_T\left(\frac{q_S}{\mu_S}\right)e^{-r/\lambda}\right]$$

where

$$\left(\Delta\frac{q}{\mu}\right)_T = \left[\left(\frac{q_A}{\mu_A}\right) - \left(\frac{q_B}{\mu_B}\right)\right]$$

and where the tests masses A and B have been prepared so that $m_A = m_B = m_T$ (where any differences will be inconsequential given the functional relationships involved).

The differential force involved therefore will be the partial derivative of this potential:

$$F = -\frac{\partial}{\partial r}V\Delta(r) = G\,\frac{m_T m_S}{r^2}\left[\alpha\left(\Delta\frac{q}{\mu}\right)_T\left(\frac{q_S}{\mu_S}\right)\left(1 + \frac{r}{\lambda}\right)e^{-r/\lambda}\right].$$

But since $F = m_T\Delta a$, it follows that

$$\alpha = \frac{\Delta a}{G\,\frac{m_S}{r^2}\left[\left(\Delta\frac{q}{\mu}\right)_T\left(\frac{q_S}{\mu_S}\right)\left(1 + \frac{r}{\lambda}\right)e^{-r/\lambda}\right]}.$$

Assume now that the 'charge' candidates have been decided upon, and that tests masses have been chosen for their ability to reveal gravitational differences (given their 'charges'), and finally that an experimental result for the difference in accelerations of test masses has been obtained. Given all this, the above expression specifies α as a function of λ. One can then solve for given values of λ to determine the corresponding strength of the non-Newtonian gravity present—assuming, of course, that the residual difference from zero in the experimental result represents more than just systematic error. The resulting solution set therefore determines a limit on how large α can be for given assumed values of λ. For a glimpse of where this is going, take a quick look at figure 5.10. Here, the area above the curve of the solution set defines where non-Newtonian forces are not present, while the area below the curve may be occupied by either unaccounted for systematic error or by some form of non-Newtonian gravitation. Thus construed, the experimental venture is to squeeze the graph ever more closely to the zero axis—or if you're an enthusiast of non-Newtonian violations of WEP—to show that it can only be squeezed so far.

The above process is therefore appropriately characterized as a *probe* or *method of discovery* precisely because it *reveals* the available real estate that may and may not be occupied by non-Newtonian gravitational effects. That's the basic idea. But there's a complication that we've glossed over that makes the process rather difficult in actual practice. As the Yukawa potential reveals, one needs to know the mass and location of the source of gravitational attraction. But since we don't live in a world

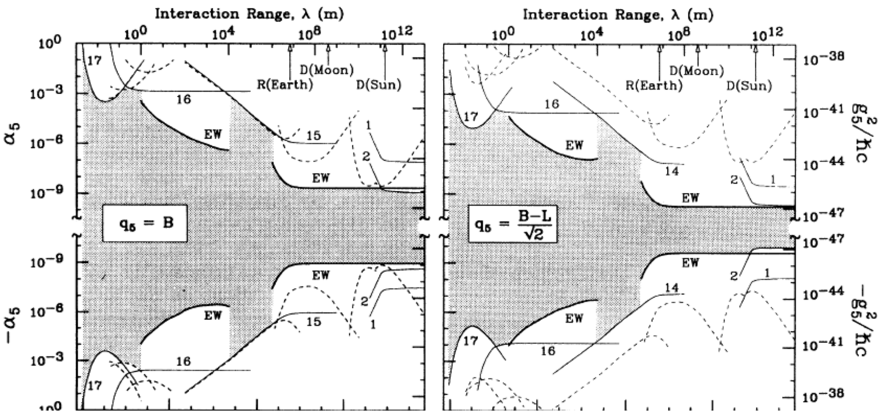

Figure 5.10. 95% confidence limits on α vs. λ for charges B and (B-L)/$\sqrt{2}$, where the heavy EW curves are as determined in the 1994 Eöt-Wash replication and the other curves are from other tests. Source: Su *et al* (1994).

of calculationally convenient point sources, this means that considerable effort is required to create a workable and accurate specification of gravitational sources of finite size and their likely non-uniform composition. Briefly stated, what one needs to do is to replace the point source attractor mass m_A with an integral taken over all the efficacious point sources in a way that takes into account actual density and distribution[20].

Since the experimental literature dealing with probing for non-Newtonian effects is quite extensive, we shall focus attention on the efforts, after the demise of the Fifth Force, of the Eöt-Wash group to both test WEP and probe the possibilities for new forms of non-Newtonian interaction[21]. In addition to being a quintessential example of 'measuring nothing, repeatedly', the replications by the Eöt-Wash group illustrate an important feature of experimental practice, what has been referred to as 'instrumental loyalty' and the 'recycling of expertise'. Given the complexity and highly developed nature of contemporary experimental practice, such loyalty and recycling is a reasonable experimental strategy—though, of course, just one among many. Moreover, as will become evident the expertise and sophistication of the Eöt-Wash group increased with each new iteration of their torsional balance experiment.

The first of the replications reported by the Eöt-Wash group (Adelberger 1990, Su 1994) were motivated by similar concerns and employed, with a few exceptions, essentially the same experimental procedure and apparatus. The motivating challenge was to achieve high accuracy with respect to forces with ranges less than the

[20] For the basics of integration over extended sources, see Fischbach and Talmadge (1999, pp 32–41); and for its incorporation in the Yukawa potential see Wagner (2012b, pp 32–3).

[21] For a concise review of much of the experimentation after the demise of the Fifth Force, see Franklin and Fischbach (2016, pp 115–44) and for yet more, see Speake and Will (2012) which is an introduction to a special issue of *Classical and Quantum Gravitation* on such experimentation.

distance to the Sun—especially short-range forces that might be expected from interactions with bosons[22].

The most stringent limits on equivalence-principle violating macroscopic forces have come from the Roll–Krotkov–Dicke and Braginsky–Panov measurements of the differential acceleration of test bodies toward the Sun. These extremely precise limits, however, do not provide strong constraints on interactions of bosons, because such bosons would produce forces with ranges less than the distance to the Sun (Adelberger *et al* 1990, p 3268).

The solar experiments, however, had two experimental advantages when it came to accuracy. First, the torsional balance remained in place and thereby eliminated systemic errors that would have been caused by its rotation. Second, '[t]he great distance to the Sun effectively eliminates problems from gravity gradients and magnetism' (Su *et al* 1994, p 3614). So the problem for the Eöt-Wash group was to achieve similar precision but with respect to short range forces, and here the solution was twofold. First, to introduce accuracy improving refinements into the type of apparatus that it had earlier used in the 1987 experimental test of the Fifth Force hypothesis. Second, to test not only with respect to the Earth as a gravitational source as but also with respect to 'a massive *laboratory source*' (Adelberger *et al* 1990, p 3268). We'll discuss the use of the 'massive laboratory force' later when we consider the Eöt-Wash 1997 and 2000 replications of that aspect of the 1990 experiment.

One feature of the 1987 experiment that was retained (and would continue to be retained throughout) was 'that all test bodies [had] to be 'identical on the outside', i.e. to have the same outside dimensions regardless of their density, and to have the same chemical surface' (Adelberger 1990, pp 3269–70). As we noted earlier, this was an updated version of Newton's similar constraint on his pendulum experiment. Consistent with this overriding requirement, there were significant improvements in both the apparatus, experimental procedure and data analysis in both the 1990 and 1994 experiments. Readers are encouraged to take a quick look at the many refinements involved in these and the later Eöt-Wash replications in order to develop an appreciation for the extensive—and often overwhelming—*accumulated expertise*. In the case of these experiments 'instrumental loyalty' and the 'recycling of expertise' were essential in achieving the *increasingly* more accurate and extensive results. Since our aim is to explicate the sense in which these experiments were used as 'probes for new interactions' we shall forgo—however inviting—a review of these many improvements[23].

The starting point of any such probe was a determination of the differential accelerations of the test masses. The 'charges' suspected of being efficacious were B, L, B—L and 3B + L. Assuming this set of suspects, the test body pairs used were

[22] The Eöt-Wash experiments to be discussed here also involved determinations of the differential acceleration toward the Sun and other astronomical entities, as well as applications in quantum physics. But in the interests of maintaining a reasonable focus on the basics of the use of the torsional balance to probe the limits of non-Newtonian gravitation, we shall pass over these determinations and applications.

[23] For an extensive review, see Wagner (2012b, pp 58–154).

beryllium–aluminum (Be-Al) and beryllium-copper (Be-Cu) in order to maximize these charge differences and to thereby distinguish their test masses from those used in the solar experiments. See Adelberger (1990, p 3283) and Su (1994, p 3615).

The differential acceleration results (south) from the 1990 experiment (Adelberger 1990, p 3287) were:

$$m_i/m_g(\mathrm{Cu}) - m_i/m_g(\mathrm{Be}) = (0.2 \pm 1.0) \times 10^{-11}$$

$$m_i/m_g(\mathrm{Al}) - m_i/m_g(\mathrm{Be}) = (0.5 \pm 1.3) \times 10^{-11}.$$

The 1994 replication with its additional refinements *upped the accuracy by an order of magnitude* where the horizontal differential accelerations were reported as:

$$\mathrm{Be{-}Al:}\ [(-2.3 \pm 4.6)\hat{e} + (-0.3 \pm 4.6)\hat{n}] \times 10^{-12}\ \mathrm{cm\ s^{-2}}$$

$$\mathrm{Be{-}Cu:}\ [(-3.6 \pm 4.1)\hat{e} + (-3.2 \pm 4.1)]\hat{n} \times 10^{-12}\ \mathrm{cm\ s^{-2}}$$

where '$\hat{e}$ and $\hat{n}$ are unit vectors pointing east and north, respectively' (Su 1994, pp 3628–9).

The next order of business was to introduce (as described above) the Yukawa potential into the mix and using the above values for the differential accelerations to determine the limits on α (the strength of the non-Newtonian gravitation) for given values of λ (the interaction range) and thereby to probe the limits of non-Newtonian forces. As noted earlier, a significant problem for such probing is how to take into account the distribution and nature of attractor masses both nearby and distant. For $\lambda \leqslant 20$ km, the following procedure was used:

> The [gravitational source integrals] were computed using measurements of the laboratory building and detailed topographic maps of the surrounding territory out to a radius of 40 km. Conventional topographic maps gave the elevation of the soil (assumed to have a density of 2.2 g cm^{-1}) and the depth of the surrounding bodies of water. The depth of the soil–rock interface was obtained from U.S. Geological Survey (USGS) data; the rocks were assumed to have a density of 2.7 g cm^{-2}. (Adelberger 1990, p 3283)

An analysis involving the range 20 km $< \lambda <$ 1000 km, however, was not even attempted because:

> … at this length scale uncertainties from possible deeper lying density contrasts have a relatively large effect on the horizontal components of [the gravitational source gradients] to which our device is sensitive. (Adelberger 1990, p 3283)

For the range $\lambda \geqslant 1000$ km, the procedure was to compute the gravitational source 'integrals using a layered, ellipsoidal model of the Earth which assumes that the Earth is in isostatic equilibrium under gravitational and centrifugal forces' where these computations made use of available density profiles and chemical composition data (Adelberger 1990, p 3283).

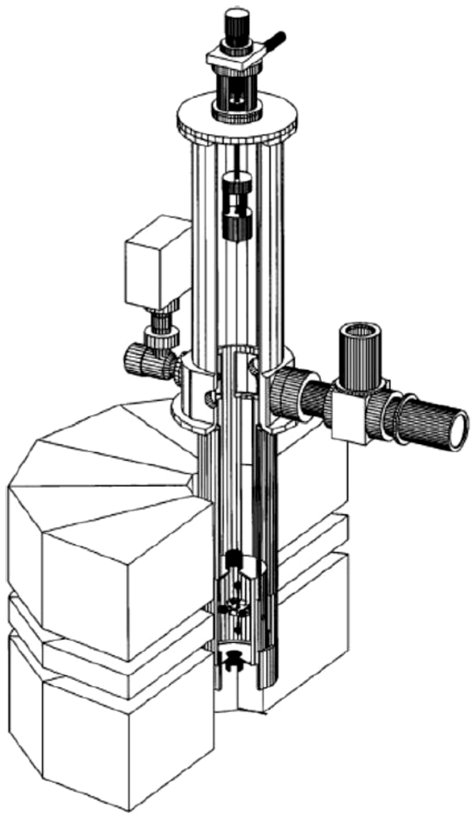

Figure 5.11. Schematic view of the Röt-Wash apparatus with the turntable omitted for clarity. Source: Gundlach (1997).

This approach, including acceptance of the λ gap from 20 to 1000km, was also adopted in the 1994 replication but with a more developed and explicit determination of the 'uncertainties' involved (see Su (1994, pp 3629–30)). The results for the variation of α with respect to λ are graphically displayed in figure 5.10[24]. The group's enthusiasm for their results was however tempered by the observation that:

> [t]he 'gap' in our results between $\lambda = 10$ km and $\lambda = 1000$ km occurs because it is difficult to model Earth's mass distribution with sufficient accuracy on length scales for which Earth cannot be approximated as a fluid in equilibrium under gravitational and centrifugal forces; nor can it be modeled simply in terms of the surface topography and bedrock profiles. (Su 1994, p 3629)

[24] The mirror image in the bottom half reflects the fact that the experimental data does not determine whether the non-Newtonian force is attractive or repulsive. Later graphs would eliminate the redundancy by using the absolute value of the strength of the Yukawa non-Newtonian interaction, i.e. $|\alpha|$.

As noted above, the Eöt-Wash group supplemented their 1990 experiment with a separate experimental determination using a 'a massive *laboratory source*' (Adelberger 1990, p 3268). The 'motivation' for using such a source was straightforward:

> … all results obtained with terrestrial sources are fairly insensitive to an interaction whose charge is proportional to B—2L = N—Z [neutrons—protons], simply because such sources contain essentially equal numbers of neutrons and protons and are effectively neutral. To remove this 'blind spot' … we conducted an experiment using a laboratory source having a substantial neutron excess. (Adelberger 1990, p 3288)

Here, the laboratory source consisted of semi-circular set of lead (Pb) blocks, with an inner 215 kg set and an outer 1080 kg set of blocks. This source was stationary and, as usual, the torsion pendulum rotated during the course of an experimental run, the blocks could be moved 180° and the experiment repeated in order to reveal and take into account asymmetrical uncertainties (Adelberger 1990, p 3289). The general methodology of using a massive laboratory source in order to overcome an insensitivity problem caused by the 'effectually neutral' character of terrestrial sources was greatly expanded in Gundlach *et al* (1997) and Smith *et al* (2000). In these replications, the lead blocks were replaced with 2620 kg of depleted uranium (^{238}U). But unlike the 1990 experiment, the nearly three tons of depleted uranium were now made to rotate around a stationary torsional balance containing 2 Cu and 2 Pb test masses (Gundlach 1997, p 2523, Smith 2000). See figure 5.11 for a representation of this new experimental arrangement fancifully dubbed the 'Röt-Wash instrument' by its creators.

The final differential acceleration result of the increasingly refined 1990, 1997 and 2000 sequence of experiments with 'massive laboratory' attractors was another order of magnitude improvement. As colorfully described:

> This work has reached a differential acceleration sensitivity of $\sim 3 \times 10^{-13}$ cm s^{-2}. If an object, initially at rest, had been given this acceleration approximately 2500 years ago and that acceleration had been maintained to this day, the object would now be moving as fast as the end of the minute hand on a standard wall clock. (Smith 2000, pp 1–19)

More soberly stated, the 'final value for the differential acceleration of Cu and Pb toward ^{238}U' (Smith 2000, pp 1–16) was:

$$\Delta a \equiv a_{Cu} - a_{Pb} = (1.0 \pm 2.6 \pm 0.9) \times 10^{-13}\ \text{cm s}^{-2}.$$

Combining this result with the Yukawa potential yields the restrictions presented in figure 5.12 on the real estate available for occupancy by non-Newtonian forces. As can be seen there is considerable improvement over the short range, as well as with respect to the 'gap' in the result of the group's 1994 experiment.

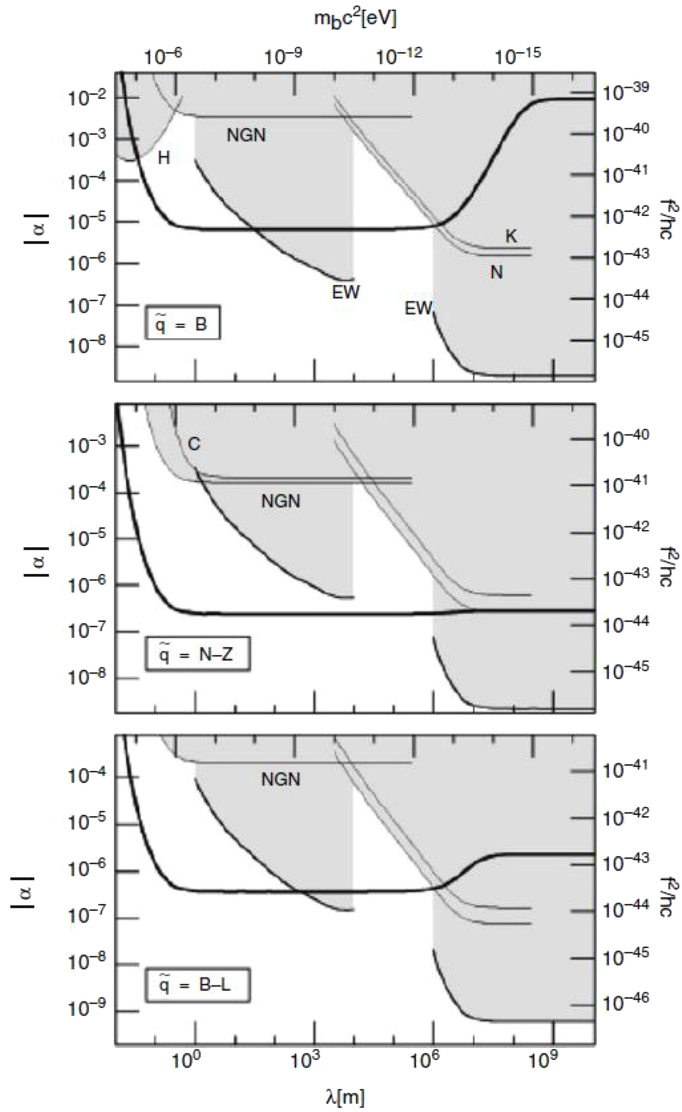

Figure 5.12. 95% confidence limits on |α| vs λ for charges B, N-Z, and B-L, where the heavy EW curves are as determined in the 2000 replication and the other curves are from other tests including the Eöt-Wash 1994 replication (denoted EW). Source: Smith (1999).

In Schlamminger (2008), the Eöt-Wash group returned to using a rotating, but greatly improved, torsional balance with the Earth and astronomical entities as gravitational attractors. There were many refinements including a newly created balance that carried four beryllium (Be) and four titanium (Ti) test masses that 'were chosen primarily to maximize the difference in baryon number (B/μ is 0.99868 for Be and 1.001077 for Ti), and secondly for experimental reasons, such as densities,

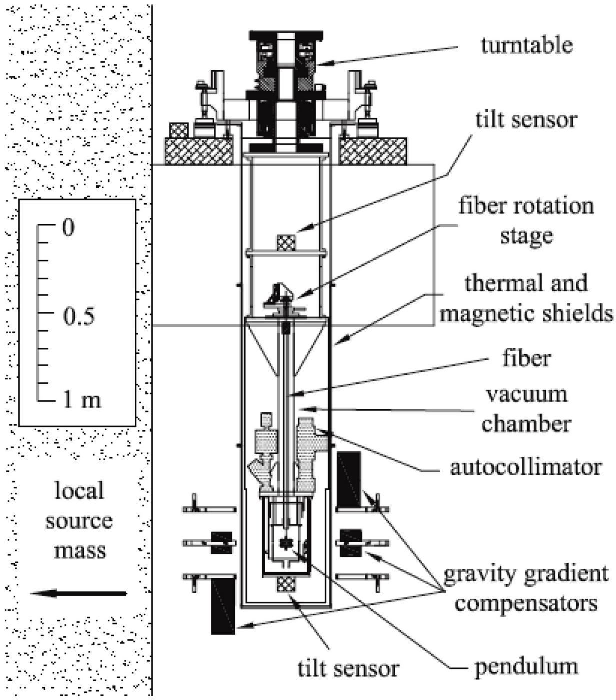

Figure 5.13. The eight test-body pendulum. Source: Schlamminger (2008).

magnetic properties, and machinability' (Schlamminger 2008) (see figure 5.13). In 2012, the Eöt-Wash group presented the results for an additional pair of test materials, aluminium and beryllium. As was the case with the 2008 pairing of titanium and beryllium, the Al-Be pairing was selected for its 'scientific impact and for practical concerns such as mechanical stability and freedom from magnetic impurities' (Wagner 2012a, p 7).

Since the test masses were removable, different configurations could be employed to test for asymmetric effects. Other disturbing causes such as tilt, temperature gradients and magnetic effects were 'deliberately exaggerated' in order to determine a measure of the systematic uncertainty involved (Schlamminger 2008, Wagner 2012a).

The final values given for the differential acceleration of the test masses indicated yet again another increase of an order of magnitude in accuracy (Wagner 2012a):

		Be-Ti	Be-Al
Δa_N	(10^{-15} m s^{-2})	0.6 ± 3.1	-1.2 ± 2.2
Δa_w	(10^{-15} m s^{-2})	-2.5 ± 3.1	0.2 ± 2.2

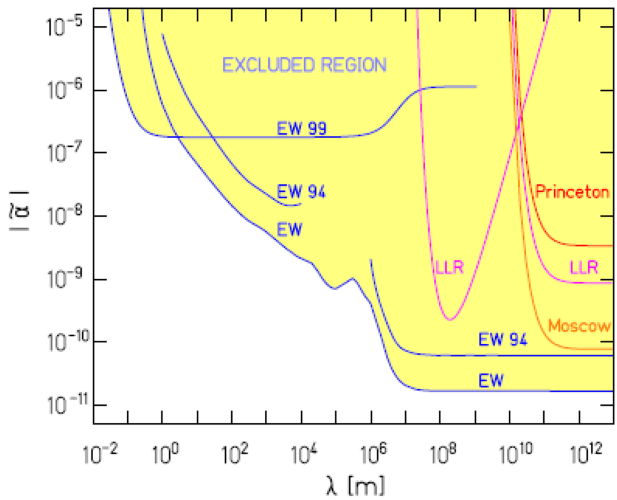

Figure 5.14. The heavy EW curve represents the 95% confidence limits on $|\alpha|$ vs λ for charge B-L as determined in the 2008/2012 replications. EW 94 and EW 99 represent the results of the Eöt-Wash replications of 1994 and 1999, while the other curves are from other experimental determinations. Source: Wagner (2012a).

In addition to the newly designed torsion balance and its myriad of refinements, the 2008 iteration by the Eöt-Wash group is noteworthy because a concentrated effort was made to develop complex density and composition models that could reliably be used to close the λ 'gap' (10 km $< \lambda <$ 1000 km) that existed in the earlier 1990 and 1994 probes[25].

The closing of this gap, and the increased precision of the values for the differential accelerations of the test mass pairs, led via the application of the Yukawa potential, to correspondingly more restrictive limits on the possible strength and range of interaction of non-Newtonian forces. See figure 5.14 which graphically, and dramatically, presents the *continual improvement* of the experimental results of the Eöt-Wash group as well as the limitations derived by other experimental efforts.

We'll close this section with a brief comment about the identification of and correction for systematic effects. By 2008, the Eöt-Wash group had become increasingly more adept at both identifying and then quantifying systematic effects. That is, at the *conversion* of *systematic uncertainty* to *quantified systematic effect*. The methodology employed by the group was a refinement of that in place since Newton's quantification of the likely effect of an hypothesized ether. In short, as noted above, to 'deliberately exaggerate' the size of the disturbing cause (Schlamminger 2008), determine the effect of the exaggeration on the quantity to be ultimately measured (the torsional twist), and then extrapolate to the likely effect on the experimental determination, i.e. the raw data[26]. With the size of the resultant

[25] See Schlamminger (2008) and Wagner *et al* (2012a). For an extended and comprehensive review of the development and use of these models, see Wagner (2012b, pp 28–54).

[26] Though even here there was a prior 'filtering' out of certain unavoidable free oscillations of the pendulum. See Wagner (2012a) and Wagner (2012b, pp 100–5).

Table 5.2. The raw differential accelerations between Be and Ti towards north (N) and west (W) are shown in line 1. Lines 2–5 list corrections that were applied, and the bottom line gives our corrected results. Source: Schlamminger (2008).

Differential acceleration in	$\Delta a_{N,\text{Be-Ti}}$ (10^{-15} m s^{-2})	$\Delta a_{W,\text{Be-Ti}}$ (10^{-15} m s^{-2})
Measurement and statistical uncertainty	3.3 ± 2.5	−2.4 ± 2.4
Residual gravity gradients	1.6 ± 0.2	0.3 ± 1.7
Tilt	1.2 ± 0.6	−0.2 ± 0.7
Magnetic	0 ± 0.3	0 ± 0.3
Temperature gradients	0 ± 1.7	0 ± 1.7
Corrected	0.6 ± 3.1	−2.5 ± 3.5

effect (for each of the identified disturbing causes) thus determined, the experimental determination of the angle of twist could be 'corrected' so as to yield the true value.

What's somewhat surprising about the end result is that the final corrected determination of the differential acceleration in one of the apparatus orientations was only a small fraction of what it would have been if determined directly on the basis of the observed values of torsional twist. See table 5.2 where the corrected value for the northerly differential acceleration (0.6) is less than 20% of the value of the raw uncorrected data. The westerly differential acceleration (−2.5) fares better insofar as it's within five percent of the raw uncorrected data. But the northerly acceleration is around four times closer to zero—though both values are less than their estimated uncertainty.

Thus, in cases such as this, one measures 'something', after which there is additional experimental manipulation of the experimental conditions which allows for a 'correction' of the value of the 'something' measured to yield a value of the 'nothing' that would have resulted in a counterfactual world without disturbing causes. The result of this series of 'measuring nothings' is that the WEP and the General Theory of Relativity are strongly confirmed and that the probe for new interactions yielded a null result.

References

Adelberger E G, Heckel B R, Stubbs C W and Rogers W F 1991 Searches for new macroscopic forces *Annu. Rev. Nucl. Part. Sci.* **41** 269–320

Adelberger E G, Stubbs C W, Heckel B R, Su Y, Swanson H E, Smith G, Gundlach J H and Rogers W F 1990 Testing the equivalence principle in the field of the Earth: particle physics at masses below 1 μeV? *Phys. Rev.* D **42** 3267

Bennett W R 1989a Modulated-source Eötvös experiment at little goose lock *Phys. Rev. Lett.* **62** 365–68

Bizzeti P G 1986 Significance of the Eötvös method for the investigation of intermediate range forces *Il Nuovo Cimento* B **94** 80–6

Bizzeti P G 1987 Forces on a floating body: another way to search for long-range, B-dependent interactions *Moriond Workshop* ed O Fackler and T T M Van (Gif sur Yvette: Editions Frontieres), pp 591–8

Bizzeti P G, Bizzeti-Sona A M, Fazzini T, Perego A and Taccetti N 1988 New search for the 'Fifth Force' with the floating-body method: status of the Vallambrosa experiment *5th Force Neutrino Physics: 8th Moriond Workshop* (Gif sur Yvette: Editions Frontieres), pp 511–24

Bizzeti P G, Bizzeti-Sona A M, Fazzini T, Perego A and Taccetti N 1989b Search for a composition-dependent Fifth Force *Phys. Rev. Lett.* **62** 2901

Bizzeti P G, Bizzeti-Sona A M, Fazzini T, Perego A and Taccetti N 1990 Recent tests of the Vallambrosa experiment *Moriond Workshop* (Gif sur Yvette: Editions Frontieres), pp 263–8

Bizzeti P G, Bizzeti-Sona A M, Fazzini T, Perego A and Taccetti N 1989a Search for a composition dependent Fifth Force: results of the Vallombrosa experiment *Moriond Workshop* (Gif sur Yvette: Editions Frontieres)

Boynton P E, Crosby D, Ekstrom P and Szumilo A 1987 Search for an intermediate-range composition-dependent force *Phys. Rev. Lett.* **59** 1385

Braginskii V B and Panov V I 1972 Verification of the equivalence of inertial and gravitational mass *JETP Lett.* **34** 463–6

Cowsik R, Krishnan N, Tandon S N and Unnikrishnan S 1990 Strength of intermediate-range forces coupling to isospin *Phys. Rev. Lett.* **64** 336

Chu S Y and Dicke R H 1986 New force or thermal gradient in the Eötvös experiment? *Phys. Rev. Lett.* **57** 1823–4

De Rujula A 1986a Fundamental forces: are there more than four? *Nature* **323** 760–1

De Rujula A 1986b On weaker forces than gravity *Phys. Lett.* B **180** 213–20

Eckhardt D H 1986 Comment on 'Reanalysis of the Eötvös experiment' *Phys. Rev. Lett.* **57** 2868

Eötvös R, Pekár D and Fekete E 1922 Beiträge zum gesetze der proportionalität von trägheit und gravität *Annelen der Physik (Leipzig)* **68** 11–66

Fischbach E, Talmadge C and Aronson S H 1986b Response to Eckhardt *Phys. Rev. Lett.* **57** 2869

Fischbach E, Sudarsky D, Szafer A, Talmadge C and Aronson S H 1986 Reanalysis of the Eötvös experiment *Phys. Rev. Lett.* **56** 3–6

Fischbach E, Sudarsky D, Szafer A, Talmadge C and Aronson S H 1988 Long-range forces and the Eötvös experiment *Ann. Phys.* **183** 1–89

Fischbach E and Talmadge C L 1999 *The Search for Non-Newtonian Gravity* (Heidelberg: Springer)

Fitch V L, Isaila M V and Palmer M A 1988 Limits on the existence of a material-dependent intermediate-range force *Phys. Rev. Lett.* **60** 1801

Franklin A and Fischbach E 2016 *The Rise and Fall of the Fifth Force* (Heidelberg: Springer)

Fujii Y 1971 Dilatonal possible non-Newtonian gravity *Nature (Phys. Sci.)* **234** 5–7

Gundlach J H, Smith G L, Adelberger E G, Heckel B R and Swanson H E 1997 Short-range test of the equivalence principle *Phys. Rev. Lett.* **78** 2523

Hall A M and Armbruster H 1991 *Progress in High Energy Physics* (Amsterdam: North Holland), pp 325–39

Heckel B R 1989 Experimental bounds on interactions mediated by ultralow-mass Bosons *Phys. Rev. Lett.* **63** 2705–08

Keyser P T 1989 Forces on the Thieberger accelerometer *Phys. Rev. Lett.* **62** 2332

Kuroda K and Mio N 1990 Limits on a possible composition-dependent force by a Galilean experiment *Phys. Rev.* D **42** 3903–7

Milgrom M 1986 On the use of Eötvös-type experiments to detect medium-range forces *Nucl. Phys.* B **277** 509–12

Nelson P G, Graham D M and Newman R D 1990 Search for an intermediate-range composition-dependent force coupling to *N-Z* *Phys. Rev.* D **42** 963

Neufeld D A 1986 Upper limit on any intermediate-range force associated with Baryon number *Phys. Rev. Lett.* **56** 2344–6

Newton I 1999 *The Principia, Mathematical Principles of Natural Philosophy* (Berkeley, CA: University of California Press)

Niebauer T M, McHugh M P and Faller J E 1987 Galilean test for the Fifth Force *Phys. Rev. Lett.* **59** 609–12

Raab F J *et al* 1987 Search for an intermediate-range interaction: results of the Eot-Wash I experiment *New and Exotic Phenomena: Seventh Moriond Workshop* ed O Fackler and J Tran Than Van (Gif sur Yvette: Editions Frontieres)

Roll P G, Krotkov R and Dicke R H 1964 The equivalence of inertial and passive gravitational mass *Ann. Phys.* **26** 442–517

Schlamminger S, Choi K Y, Wagner T A, Gundlach J H and Adelberger E G 2008 Test of the equivalence principle using a rotating torsion balance *Phys. Rev. Lett.* **100** 041101

Smith G L, Hoyle C D, Gundlach J H, Adelberger E G, Heckel B R and Swanson H E 2000 Short-range tests of the equivalence principle *Phys. Rev.* D **61** 022001

Speake C C and Quinn T J 1988 Search for a short-range, isospin coupling component of the Fifth Force with the use of a beam balance *Phys. Rev. Lett.* **61** 1340–3

Speake C C and Will C M 2012 Tests of the weak equivalence principle *Class. Quant. Grav.* **29** 180301

Stubbs C W, Adelberger E G, Raab F J, Gundlach J H, Heckel B R, McMurry K D, Swanson H E and Watanabe R 1987 Search for an intermediate-range interaction *Phys. Rev. Lett.* **58** 1070–3

Stubbs C W 1989b Limits on composition-dependent interactions using a laboratory source: is there a ‘Fifth Force’ coupled to isospin? *Phys. Rev. Lett.* **62** 609–12

Su Y, Heckel B R, Adelberger E G, Gundlach J H, Harris M, Smith G L and Swanson H E 1994 New tests of the universality of free fall *Phys. Rev.* D **50** 3614–36

Talmadge C, Aronson S H and Fischbach E 1986 Effects of local mass anomalies in Eötvös type experiments *Moriond Workshop* (Gif sur Yvette: Editions Frontieres)

Thieberger P 1986 Hypercharge fields and Eötvös-type experiments *Phys. Rev. Lett.* **56** 2347–9

Thieberger P 1987a Search for a new force *New and Exotic Phenomena: Seventh Moriond Workshop* (Gif sur Yvette: Editions Frontieres), pp 579–89

Thieberger P 1987b Search for a substance-dependent force with a new differential accelerometer *Phys. Rev. Lett.* **58** 1066–9

Thieberger P 1989 Thieberger replies *Phys. Rev. Lett.* **62** 2333

Wagner T A 2012b Rotating torsion balance test of the weak equivalence principle *Dissertation* University of Washington

Wagner T A, Schlamminger S, Gundlach J H and Adelberger E G 2012a Torsion-balance tests of the weak equivalence principle *Class. Quant. Grav.* **29** 184002–15