XIII

Three Case Studies

In this chapter, three case studies set outside of the usual application of latent variable models to sets of test items are considered[1]. The purpose is to illuminate points made, often in the abstract, earlier in the chapter. The first concerns the discipline of *sabermetrics*, the mathematical and statistical analysis of baseball records. The second concerns the application of linear factor analysis to recordings of a signal from space, and the third, the application of linear factor analysis to a set of scalp eeg (electroencephalogram) recordings.

1. Sabermetrics and hitting ability

Sabermetrics is the mathematical and statistical study of baseball performance data. This data includes performance indices for hitters, pitchers, and teams. Among other indices, hitters are described by the numbers of singles, doubles, triples, and home runs that they produce, and walks that they draw. A chief aim of the sabermetrician is to adjudicate on the hitting abilities of baseball players, and, more generally, to discover and investigate the determinants of hitting ability. The scientific context of this endeavour can be described as follows.

There is a concept *hitting ability*, whose correct employments are, as with every concept, fixed by linguistic rules. Linguistic rules are layed down by humans. Humans invented the concept *hitting ability* to denote certain phenomena. Ability terms are ascribed to others on the basis of their behaviours. The possessor of a particular ability may exercise his ability, or choose not to. When former major league all-star Kirby Puckett decided to "hang em' up", he chose to cease exercising his hitting ability in a professional context. As with so many ability and capacity terms, the grounds of correct ascription of the term *hitting ability* are relatively loose. There are many criteria that can legitimately be cited in support of a claim that player A is a poor hitter or player B is superb. The concept is not correctly ascribed on the basis of necessary and sufficient conditions. The relatively messy grammars of ability terms does not mean that they are *flawed*, for these terms satisfy our everyday needs to communicate, which is precisely their purpose. It does suggest, however, that they might well turn out to be problematic as conceptual foundations for scientific work. For fruitful scientific work on, say, φ-entities, requires that the scientist grasp what are and are not φ-entities, to grasp what are and are not φ-entities is to grasp what concept "φ" denotes, and to grasp what concept "φ" denotes is to understand its grammar. However, unlike the relatively simple grammars of technical concepts, the grammars of ordinary language ability terms do not support precise denotative relations.

To correctly claim that "certain baseball players have a greater ability to hit than others" or that "player A is a talented hitter" is not to make an inference, but rather to ascribe particular concepts in accord with the rules that fix their correct employments. One ascribes terms such as *good hitter*, *bad hitter*, *talented hitter*, etc., to baseball players on the basis of their hitting behaviours, and baseball players differ in their hitting behaviours. Grammar (rules of language)

---

1 The author thanks Professor James Steiger for inventing the first and last case study for the puropse of clarifying the ideas discussed in this book.

settles whether there are grounds for distinguishing between baseball players on the basis of their hitting abilities, and the concept *hitting ability* (related terms and cognates) does in fact admit of predication by terms such as *supreme*, *inadequate*, *good*, *bad*, just as colour terms admit of predication by modifiers such as *rich*, *pale*, *deep*, etc. We are not in need of something *in* baseball players, or anywhere else, to warrant the judgments we make about their hitting abilities.

For various practical reasons (chiefly, the desire to proclaim a "batting champion" at the end of each baseball season), more clear-cut grounds of ascription of the concept *hitting ability* have been required, and this has led to the taking of batting average (the average number of hits per at bat) as a technical definition of *hitting ability*. The installation of this technical definition has not pleased everyone, as adherence to it prohibits ascription of the concept *hitting ability* to a baseball player on *other* grounds that are part of the loose, ordinary language sense of the concept (e.g., a player's overall run production). Thus, as in every legitimate science, there has ensued a great deal of argumentation amongst sabermetricians over how the concept *hitting ability* should be defined. This marks the sabermetrician's study of hitting ability as at the pre-empirical stage of science.

There is nothing "out there" in nature (unbeknownst, or only partially known, to humans) called hitting ability. Sabermetricians will not be discovering what the term *hitting ability* "really" means. Nor are they currently moving toward an understanding of what hitting ability "truly" is. If things continue to proceed in a scientific manner, sabermetricians will likely come to an agreement as to how they would like to define the concept *hitting ability*. To reach this point will likely take a great deal of argumentation and analysis of possible choices. For example, the unsatisfactoriness of taking hitting ability to be batting average is believed by some to be shown by the fact that a player who averages one home run per at bat will have to be said to have the same hitting ability as a player who averages one single per at bat, despite the fact that the former hitter will produce many more runs than the latter. When sabermetricians settle on a technical definition of the concept *hitting ability*, they will then have reached agreement on the grounds of correct ascription of a technical sense of the term to baseball players. Only then will it be clear what must be studied in an empirical study of hitting ability. *Pace* the construct validators, sidestepping this conceptual work in favour of empirical research will not lead to the accumulation of evidence bearing on what is *meant* by the concept *hitting ability*. It is humans who *assign* senses to their scientific terms.

Regardless of the technical definition of *hitting ability* that will eventually be adopted by sabermetricians, if hitting ability is truly an ability, it will be ascribed to baseball players on the basis of their hitting behaviours. While the causes of, e.g., sublime hitting ability, may well turn out to be unobservable, the adjective *unobservable* will not be correctly ascribable to what the concept *hitting ability* denotes. Hitting ability is a capacity and, hence, is neither observable nor unobservable, but, rather, exercised or not. What is manifested and produced in the exercise of a baseball player's capacity to hit, i.e., singles, doubles, home runs, etc. (these criteria for ascription of the term), is certainly observable. What made Kirby Puckett such a great hitter may well have been a unique neuro-muscular makeup, but the term *great hitter* was rightly ascribed to Kirby on the basis of what he *did* on the baseball field. If sabermetricians come to misrepresent their conceptual difficulties (the fact that they are currently unable to settle on a satisfactory technical definition of the term *hitting ability*) to be the result of the unobservability of some *thing* denoted by the concept *hitting ability*, they will then consign their discipline to the status of a pseudo-science. Only constituents of natural reality that can potentially reflect light (e.g., material entities) can be coherently placed on the dimension of unobservable to observable, but capacity terms such as hitting ability do not denote material entities, and, hence, do not denote material entities that happen not to reflect light. From at least Spearman's time, social scientists have mistaken the complexities of the grammars of psychological concepts as implying that they have an unobservability problem on their hands. This misdiagnosis has led them to invent various "solutions", notably, the Central Account and construct validation theory, that have drawn the discipline into the domain of pseudo-science.

There are at least two distinct senses in which hitting ability can be said to be complex:

i. The ordinary language sense of the concept *hitting ability* has a complex grammar. That is, the rules that fix the correct employments of the concept are not of the simple necessary and sufficient condition type. There are multiple and diffuse criteria for the ascription of the term. The grounds for correctly ascribing the concept to a baseball player might be said to be messy. This grammatical messiness can create a feeling that the concept is ineffable, and that this has something to do with unobservability. However, no concept is ineffable (else, it would not be a concept). If one adopts the technical homynym that takes batting average to be hitting ability, then one side-steps the problem of a complex grammar, but, on the other hand, is restricted to studying a phenomenon that is different than that which was originally of interest.

ii. When sabermetricians succeed in clearing up their equivocation over the concept *hitting ability*, they will have succeeded in clarifying the phenomena that they must study in a study of the hitting abilities of baseball players. They will, only then, have something to empirically investigate in a study of hitting ability. This something will undoubtedly turn out to be empirically complex. Empirical investigation will, undoubtedly, involve the complex issue of the causes of superior hitting ability.

Psychologists often misportray (ii) as a case of "conceptual conservativism", and are assured by the tenets of construct validity that "every science gets on with empirical study long before the "meaning" of the phenomena under study is fully grasped." But this reassuring tenet is a product of construct validation's conflation of conceptual and empirical issues. Of course it is true that sciences get on with studying empirical phenomena long before phenomena of interest are fully understood. In fact, sciences get on with such study just *because* phenomena are not understood. This trivial observation has no bearing whatever on the temporal priority of conceptual clarification in the study of the empirical natures of, say, ϕ-phenomena. The conceptual features of concept "ϕ" must be clarified prior to the fruitful scientific study of ϕ-phenomena because to study ϕ-phenomena is to study that which is denoted by concept "ϕ". The scientist who is going to generate facts *about* ϕ-phenomena must, among other things, be able to identify ϕ-phenomena in nature, and distinguish it from other types of phenomena. The conceptual features of concept "ϕ" are conceptual properties and are logically independent of the empirical features of the phenomena concept "ϕ" denotes (just as the grounds for ascribing the concept *bachelor* to an individual are logically distinct from properties *of* bachelors. To discover that bachelors have characteristic t would *require* the capacity to correctly identify bachelors and distinguish such individuals from those who are not bachelors, which is equivalent to possessing the grounds for ascribing the concept *bachelor* to individuals).

Consider the technical definition of *hitting ability* as batting average. Given that many are dissatisfied with this choice, can it rightly be said that batting average is an "imperfect measure" of hitting ability? Do the deliberations that sabermetricians have been engaging in over a satisfactory definition of the concept *hitting ability* represent steps in the improvement of the measurement of hitting ability? *A* can only be called an imperfect measure of *B* (and what is really meant here is that *A* is an imperfect *way of taking measurements* of *B*) if there is a *sense* to the notion of measuring *B*. And, as it stands, there is no sense to the notion of correctly measuring hitting ability because, as it stands, there is rampant equivocation amongst the sabermetricians over how they would like to employ the term *hitting ability*. If anything, sabermetricians are making progress toward settling how they would like to employ the concept *hitting ability*, and this brand of progress is entirely different than progress in the measuring of something. When the sabermetricians eventually settle their conceptual issues and come to speak unambiguously about hitting ability, there will then arise legitimate measurement problems and the potential for refinements in measurement. Physics, for example, has made progress in the measurement of length. Whereas, the meaning of *one metre* was once fixed by canonical sample (the standard metre), it is now fixed by canonical recipe (Def'n: *one metre*: The length of the path traveled by light in a vacuum in 1/299,792,458 of a second). Similarly, there have been enormous refinements in the measurement of weight through, e.g., the invention of atomic scales. But such refinements are only possible given a settled conceptual foundation.

Now, consider the case in which one took the path of the latent variable modeler under the influence of the Central Account. One would portray hitting ability as an unobservable attribute (cause) and the array of performance indices (number of singles, doubles, etc.) as observable "indicators" of this attribute (cause). One would bypass the conceptual clarifications that are a precondition of fruitful empirical work, and would, instead, attempt to access the "unobservable" hitting ability by, say, factor analyzing the performance indices. If the indices turned out to be unidimensional in a linear factor analytic sense, one would take this as evidence that hitting ability had been detected. Perhaps this factor would be labeled "general hitting ability" and some composite of the indices formed in an attempt to estimate players' true values on this "underlying" attribute.

This approach would consign the study of hitting ability to the status of pseudoscience.

i. It would constitute a misdiagnosis of a conceptual problem (equivocation over the grounds of ascription of the concept *hitting ability*) as an empirical problem (the unobservability of a phenomenon of interest). But the resolution of the conceptual problem was a *precondition* for the fruitful scientific study of hitting ability, for only when an acceptable technical usage of the concept *hitting ability* has been settled on will the *target of* empirical investigation, the hitting abilities of baseball players, be known. What is meant by the concept *hitting ability* will not become clear by examining natural reality, for the phenomena of hitting ability is simply that which is denoted by the concept *hitting ability*. And it is humans that fix, through the laying down of linguistic rules, what concepts do and do not denote. Hence, the intial misdiagnosis would guarantee the production of results whose interpretations were inherently ambiguous.

ii. Far from having settled on a sense of the concept *hitting ability* that denotes a perceptually unobservable constituent of natural reality, the latter for which one could set about building a detector, there exists rampant equivocation over the concept. Thus, there can exist no perceptually unobervable referent of the concept *hitting ability* whose detection might be a possibility.

iii. Without a perceptually unobervable entity whose detection might be the subject of scientific work, a bogus detection scenario is invented with a set of equations offered up as a tool of detection.

iv. If a new technical sense of the concept *hitting ability* were to be invented and this sense warranted ascription of the concept to some perceptually unobservable constituent of natural reality (e.g., some brain centre), then latent variable models would be of no relevance because constituents of natural reality are detected through their physical properties (e.g., the electrical or magnetic fields they emit, the reactions they cause). Latent variable models, on the other hand, are not detectors of anything. Hence, detectors built to detect the physical properties of the referent of the concept *hitting ability* would then be of relevance.

Generally speaking, what could a linear factor analysis of a set of indices tell us about hitting ability? One can only generate findings about φ-phenomena if one can study φ-phenomena, and to do so presupposes the ability to identify φ-phenomena in nature. But the ability to identify φ-phenomena in nature is equivalent to the ability to correctly ascribe a concept "φ" that denotes this phenomena, which, in turn, is equivalent to grasping the grammar of this concept. But these ingredients are precisely what are, herein, lacking. There is currently equivocation over the concept *hitting ability*, and until this equivocation is removed, it will not be clear what is meant by a study of hitting ability. What then could one say if a large set of performance indices were, say, unidimensional in the linear factor analytic sense? Certainly, one could not coherently claim that anything has been detected, nor that "the common factor is hitting ability", nor that hitting ability is "a single thing." Factor analysis is not a tool of detection and, hence, its employment cannot support claims of detection. Nor is the researcher in any position to make the case that the concept *hitting ability* is rightly ascribed to the scores comprising the distribution of one of the common factors to a set of indices, a case that would have to be made to support the claim that "the common factor is hitting ability." Whether hitting ability is a single thing or many things is determined by the rules that fix the correct employments of the concept *hitting ability*, and, in particular, whether these rules allow for the correct ascription of the concept to different types of thing (i.e., in a manner akin to the way in which the concept *game* is ascribable to a number of very different types of thing).

On the other hand, it is certainly not the case that unidimensionality implies that the behaviour represented by the indices has but a single cause, for, once again, latent variable models are not tools of detection. Bubble chambers are tools of detection, and the presence of a certain type of track through a bubble chamber is evidence of the existence of an unobservable material entity called a subatomic particle. The bubble chamber was invented on the basis of knowledge about how such particles behave. That is, it is predicated on discoveries of correlational relationships between particle types and track types. In contrast, the unidimensionality of a set of variates was *not* something that was discovered to be the outcome of the impact of some *thing* to be detected upon the phenomena represented by the variates. The causal story of the behaviours represented by the indices is an empirical issue that latent variable models cannot pronounce upon. The causal story of the indices is currently unknown. An understanding of this story may well be *required* to explain the unidimensional structure of the indices. Eventually, research may reveal part of this causal story, laws formulated to describe this story (these laws involving, not latent variates, but rather symbols representing the particular constituents of natural reality to which the laws refer), and the covariance structure (and much else) of the indices deduced on the basis of these laws.

What if the sabermetricians settled on a technical concept of *hitting ability*, say, "φ", that denoted members of a class of phenomena φ. Then phenomena φ denoted by this technical term is the phenomena that is to be studied in a study of hitting ability. If the indices turned out to be unidimensional, then they would be so for some currently unknown reason. The unidimensional result would be in need of explanation. The linear factor model, involving latent variates rather than symbols representing some characteristic of phenomena φ, has nothing directly to do with φ. The equations of linear factor analysis involve manifest variates that represent phenomena of interest, in this case, various characteristics of a baseball player's hitting behaviour, but the latent variates contained in these equations do not represent anything in nature. True scientific work might eventually result in the formulation of laws relating properties of φ to the phenomena represented by the indices, and this knowledge might be used to eventually explain the unidimensionality result. In the meanwhile, all that can be said is that the indices correlate in a very particular way. The discovery of, say, a large positive correlation between two indices is a discovery about a feature of the joint distribution of these indices in the population of baseball players. In particular, such a correlation says something about the typology of baseball players created when they are characterized by the indices, namely that certain types of player (those high on one index, and low on the other) do not exist in large numbers. What the special linear factor analytic correlation structure says about a set of variates has to do with the replaceability of the set by constructed random variates, and is the topic of Part 3 of this book.

2. Steiger's "signal from space" example

In a commentary on the paper *Metaphor taken as math: Indeterminacy in the factor analysis model* (Maraun, 1996), Professor James Steiger (1996) offered up the following scenario:

i. It is believed that a signal may be emitted ten, and only ten, times at one hour intervals from point *X*, starting at 1:00p.m.
ii. Point *X* can never be observed directly.
iii. Receivers are constructed to measure the signal at points Y_1, Y_2, Y_3, and Y_4.
iv. The signal is jammed by a noise countersignal at each receiving point.
v. An additional signal, at point *W*, is received directly without any noise degradation.
vi. It is known from intelligence sources that:
  a. all signal and noise distributions have zero means;
  b. the signal and noise components are additive;
  c. the signal and noise components are uncorrelated over the ten recordings taken;
  d. the ten signals have a variance of exactly unity.

Steiger further specifies that:

vii. The distribution of **Y** is discrete multivariate, with only ten 4-tuples assigned positive probability, the ten 4-tuples equally likely, so that the ten recordings mirror exactly the population multivariate distribution.

He concludes from this that the signal believed to exist can be considered a random variate **X** that satisfies the definition of a common factor to **Y**, that it can be taken as a certainty that the assumptions of the factor model actually hold, and that the linear common factor model is therefore an "appropriate data generation model for these data." The aim of his analysis was to consider whether a linear factor analysis of **Y** would be capable of answering the following questions:

1. How much noise is degrading each signal?
2. What are the values of the signals?
3. Do the signals show any intelligent pattern?
4. Is there any correlation between the signal at *X* and the signal at *W*? If so, what is it?

While fascinating, this scenario has the dangerous unresolvedness that nurtures growth of the Central Account. It could be interpreted in a number of different ways, and these different possibilities must be disambiguated.

Possibility 1 (True case of signal detection/recording)

For some reason, it is believed that ten signals are to be sent to earth, one per hour starting at 1:00p.m., from a point *X* in space. The claim that "point *X* can never be observed directly" means that the signals cannot be recorded at their source. The nature of the signals is not described by Steiger, but, because the signals are coming from space, they will either be constituted of, or be carried by, some type of electromagnetic waves. Electromagnetic waves vary with respect a number of parameters, including wave length, amplitude, and frequency. They travel at 3×10^8 m/sec. in a vacuum, and, elsewhere, somewhat slower depending on the medium through which they are passing. The ten signals incoming from space might be, among many possibilities, gamma rays, something in the visible spectrum, or long radio waves. An electromagnetic wave has both a spatio/temporal point of origin and a source (the constituent of natural reality that generated it). In Steiger's scenario, it is known that the point of origin of the incoming waves is point *X* in space. Nothing is said about what generated the signal waves, nor even whether the waves have the same or different sources, and this information might not be available. If the ten signals are to be detected and recorded, it will need to be known, or guessed correctly, the type of wave of which they are constituted. Different types of wave require different types of detectors and recording instruments. Radio waves, for example, are detected by radio telescopes, while waves of the visible spectrum are optimally detected by optical telescopes. If it is unclear what type of wave the signals will be, it will be necessary to train different types of detector on space in the direction of point *X*.

Imagine that ten signals arrive at the times predicted, happen to be radio waves contained within some frequency interval, and are detected and recorded by a cluster of four radio telescopes. A particular property of these signals (perhaps their frequency or wavelength) constitutes the data. Now, detectors such as the radio telescope are built to detect *particular* classes of natural phenomena. A track detector is built to detect sub-atomic particles, while litmus tests are built to detect particular types of chemical reaction. The radio telescope is built to detect electromagnetic waves of a particular frequency range. If the four detectors each pick up ten signals occurring at one hour intervals, then these signals exist. Because, from the physics of electromagnetic waves, it is known that no electromagnetic wave can exist without it possessing a source, it can be concluded that the signals arose from one or more sources whose natures are unspecified.

Consider certain of the ingredients of such a scenario of detection. The construction and employment of a detector exploits knowledge of the reaction of some particular media to certain of the properties the thing to be detected. And the generation of such knowledge presupposes the capacity to identify in nature that which is to be detected, and distinguish it from other distinct constituents of natural reality. To possess the capacity to identify in nature that which is to be detected, say, κ-things, is to grasp the rules of correct employment of the concept "κ" that denotes κ-things. A precondition for the construction of a radio telescope is a grasp of the linguistic rules that fix the correct employments of the concept *radio wave*. For to grasp the rules that fix the correct employments of the concept *radio wave* is to understand its legitimate denotative employments, and, hence, to possess the capacity to identify radio waves in nature and distinguish them from distinct natural phenomena. The capacity to identify radio waves in nature and distinguish them from other natural phenomena is a precondition for the accumulation of knowledge, and the generation of theory, about the behaviour of radio waves, and, in particular, what happens to them when they impact a parabolic reflector of the sort used in a radio telescope.

Physics has generated extensive knowledge and theory on the electromagnetic wave/detector relationship, and this knowledge and theory is represented by mathematical equations (mathematical laws) that relate a wide variety of the properties of such signals at source to the properties of the recordings made of them by particular detectors. These equations allow for precise deductions to be made, on the basis of recordings, about the properties of signals at source, including locations of their sources, whether or not the signals recorded by different radio telescopes at some time t_o originated from the same source, and the distance that the signals themselves have travelled through space. Much is also known about the performance of various detectors and recorders, including the performance of the radio telescope in the face of various types of interference (noise). Interferometry is the theory that describes the detection performance increment possible when a number of radio telescopes are used in combination.

What could a linear factor analysis of the recordings add? Why would one need a "data generation model"? What were needed were detectors of various types of electromagnetic waves, and detectors of wave-form signals are detectors of particular natural phenomena, and, hence, are instruments (not sets of equations). Similarly, why would one represent a signal at source as a random variate (let alone a latent variate) unless one were in the position of taking realizations of the signal at source, and the generation of the signal at source were goverened by some probabilistic law? The data, in this case, were generated through the recording of wave-form signals originating with one or more physical entities, by a cluster of radio telescopes. The "intelligence" that was required here was of the kind provided by the theory and empirical findings of physics, and this intelligence allows for the *deduction* of not only the covariance structure of the recordings, but a great deal more. Professor Steiger states that "We are taking it as a certainty that the assumptions of the factor model hold" (1996, p.540). But in a true scenario of detection whatever can be taken as known is not an "assumption", but, rather, part of the *knowledge* that is the product of coherent scientific investigation. And the generation of such knowledge presupposes the capacity to identify some particular class of things to be detected. To put this another way, if we have come to *know* the features that Professor Steiger describes in this scenario (e.g., that there is a linear relationship between some property of some type of signal at source and the recordings of the signal made by some detector), then we can express this knowledge in the form of mathematical laws. And these laws will not contain latent variates (common factors), but rather symbols that represent the particular constituents of natural reality of which the laws are about (e.g., some property of the signal at source expressed in the appropriate units). The certainty present in the application of factor analysis arises not from the fact that the "assumptions" of the "model" are known facts of a detector scenario, but from the fact that these assumptions are actually prescribed *requirements* in the *construction* of latent variates.

Possibility 2 (The Central Account)

Here, the scenario is taken just as described. By the claim that "it is believed that a signal may be emitted ten times", it is meant that "it is believed that the data are "recordings" of a "signal" from an "unobserved" point *X*". No information is given as to what the "recordings" are of, hence, the nature of the "phenomena" that is implied to be the object of detection is unspecified, hence, it is unclear whether the data constitute evidence that *it* has been detected. In the previous interpretation of the scenario, the recordings were known to be recording *of* a radio wave signal by virtue of the fact that they were detected and recorded by a radio telescope. What is now absent is knowledge of the details of a detector/signal relationship that can be used to explain the recordings, and this arises in contexts in which there is not in play a concept that denotes members of a class of phenomena that are the objects of detection. Here, one speaks of "signal", "noise", "point *X*", "receivers", and "unobserved", but these terms are literally meaningless for lack of knowledge of the rules that fix their correct employments. There is no way of knowing whether the detection of anything has occurred because it is wholly unclear *what* the target of detection was. Without a target of detection, knowledge of *its* impact on some medium that is to play the role of detector cannot be generated. It is knowledge of the detector/target relationship that allows for deductions to be made about what one should expect in detection scenarios. The notion of a signal that cannot be recorded at source has a meaning, but it is far from clear what is meant by an "unobservable signal."

Could the data be recordings of signals that have a single source? Perhaps, but it is certainly not a linear factor analysis of the data that will decide the matter, nor even "give an indication" as to whether this is the case. Electromagenetic waves are particular constituents of natural reality that have particular properties and can have impacts on other constituents of natural reality. Knowledge and theory about the behaviour of a particular type of wave provides the basis for the formulation of laws, and, from these laws, deductions of further of the properties of these waves. In the factor analytic context, unidimensionality of a set of variates is not a deduced or known consequence of some particular constituent of natural reality whose detection is of interest. Hence, if a set of variates happens to be unidimensional in a linear factor analytic sense, this fact says nothing about the existence of some additional constituent of natural reality, nor, certainly, *its* functional relationship to whatever the variates happen to represent. Factor analytic unidimensionality is a property required to obtain in order that variates that satisfy the requirements for latent variate-hood under the unidimensional linear factor model can be constructed (see Part 3 of the book).

In a proper scenario of detection, what to expect from the covariance structure of a set of recordings depends upon an understanding of the physics of the signal/detector relationship. In the context of factor analysis, linearity, uncorrelatedness of "noise" and "common factor", etc., are not "assumed" in the hope that down the road scientific endeavour will verify these properties. On the contrary, there can be no way of verifying these properties, because *latent variate to **X*** does not denote a constituent of natural reality that could have these relations with the phenomena represented by **X**. The formulas that, say, describe the relationships between properties of a radio wave at source and the recordings of the wave by a radio telescope involve symbols that represent particular constituents of natural reality. The equations of factor analysis, on the other hand, contain latent variates, and latent variates do not represent constituents of natural reality whose properties are described by the factor analytic equations. Assumption talk is prescriptive in nature, outlining the properties that a constructed random variate must have in order to correctly be called a latent variate to **X**.

3. Steiger's EEG example

Let there be a single subject. Red and green lights are flashed in an alternating fashion (red, green, red, green,...) at the subject, $\frac{1}{2}N$ times each. It is believed that the flashing of these lights will cause electrical signals to be produced by a "color perception centre" (cpc), and that these signals will differ depending on the colour of the lights. Suppose it is technically impossible to insert an electrode into the brain to measure the signal directly at the cpc. It is believed that the electrical output of the cpc will be detectable on the surface of the scalp, and, to this end, electrical activity is recorded at p electrode sites following each of the N flashes. The collected data is then factor analyzed. What might such a factor analysis contribute to an understanding of the cpc? To answer this question, two very different scenarios must be distinguished.

Scenario 1:

If it is *known* that it is technically impossible to insert an electrode into the brain so as to measure the electrical output of the cpc, then it must be known that the cpc exists. It is coherent to claim "it is not known at present if there exists a cpc, but if there does, it is likely that we will not be able to insert an electrode into it", but not to make a claim of *knowledge* about something that might not exist. As are all things that can produce an electrical output, the cpc is a constituent of natural reality (in this case, a material structure). All material structures have a location, and this particular material structure is located in the brain. Because we *know* that we can't insert an electrode into the cpc to measure its electrical output, it must be the case that we know the location within the brain of the cpc (else, we would instead hypothesize the existence of a cpc). Hence, the cpc exists and its location within the brain is known to us, but, for unspecified technical reasons, it currently cannot be accessed. This is clearly a problem whose solution is going to require technical advances in neurosurgery.

As with most of the electrical activity produced within the brain, some of the activity produced by the cpc is detectable on the surface of the scalp. There exist many types of detectors that can be used to detect and record the electrical and magnetic fields that emanate from the brain. That science is in the position of constructing detectors of electrical activity follows from there being clarity wtih respect the correct employment of the term *electrical activity*. This term denotes a particular type of natural phenomenon. Thus, it is clear what is *meant* when a scientist claims that she has "detected an electrical field on the surface of the scalp." It is clear what the term denotes. The reactions of any proposed detector when in the presence of an electrical field can be studied just because electrical activity can actually be identified in nature. Given that an electrode can't be inserted into the cpc, the scientist is restricted to establishing correlations between stimuli presentation and scalp electrical activity. Often, recorded scalp electrical activity is used in conjunction with knowledge of the physics of electricity, to deduce the brain area from which such activity was likely to have originated.

Now, let us turn to the factor analysis of the electrical recordings at p scalp locations. Certainly, even if it could play that role, factor analysis would not be needed as a tool to detect, or *verify*, the existence of, the cpc. If we can non-vacuously state that "it is technically impossible to insert an electrode into the cpc", then we *know* that the cpc exists. The electrical activity produced in brain centers as a result of flashing lights is a complicated issue. Will the cpc produce exactly the same electrical response for every red flash (green flash)? It is likely that there will be a distribution of responses conditional on each of the red and green stimuli. Thus, if **X** stands for the voltage response of the cpc, it will likely be the case that, for an infinity of flashes in which the proportion of red flashes and green flashes are equal, the distribution of **X** will be a mixture of the form $.5f_{X|r} + .5f_{X|g}$, in which $f_{X|r}$ and $f_{X|g}$ are the conditional distributions of **X** given a red and green stimulus, respectively. Of course, it is research on the cpc, presumably taking place after neuroscience has overcome the earlier technical problem that prohibited measuring the electrical ouput of the cpc at source, that is the arbiter of the correctness of such a conjecture.

Will the electrical activity produced by the cpc be transmitted to the p scalp locations in a linear fashion? Will errors made in measuring the voltages recorded at scalp be uncorrelated with the voltages output by the cpc at source? These are empirical questions the answering of which will require the invention by scientists of some way of measuring the output of the cpc at source (i.e., the overcoming of the technical difficulties that currently exist). If it were the case that these obtained, then it would follow that the p scalp recordings would indeed be unidimensional in a linear factor analytic sense, and have covariance structure $\Sigma = \Psi + .25(v_r - v_g)\underline{\Lambda}\underline{\Lambda}'$, in which v_r and v_g are the mean voltages output by the cpc following a red and green pulse, respectively. What would be the structure of matrix Ψ, which represents the noise characteristics of the scalp electrodes? Only knowledge of the physics of such signal- recorder relationships could answer this question. The point is that such a covariance structure is *deduced* from knowledge and theory about particular constituents of natural reality. Whether such a deduced covariance structure squares with reality could be tested directly without recourse to latent variable modeling. If properly elaborated, it would be a posited *law* describing the relations between constituents of natural reality, and this is why the symbols it contains actually stand for particular constituents of natural reality. If science failed to overcome the technical problem that obstructs it from measuring the electrical output of the cpc at source, it would have to satisfy itself with the development of laws linking properties of colour stimuli to properties of the scalp recordings.

A factor analysis of the scalp recordings would allow us to assess the dimensionality, in a linear factor analytic sense, of our p variates. The dimensionality of a set of variates is an empirical issue. If the variates were, say, unidimensional, then this would be a result in need of explaining. Discoveries about the cpc and the consequent formulation of laws about these discoveries might well turn out to provide the basis for an explanation of the unidimensionality result. The factor analytic result itself is not *about* the cpc, but, rather, the scalp recordings. The equations of factor analysis do not describe the relationship between the cpc and the scalp recordings, nor even properties of the stimuli and the scalp recordings. Laws would have to be developed to describe such relationships. The latent variate featured in the equations of factor analysis is not a representer of some constituent of natural reality (such as the cpc or properties of the stimuli), but, instead, a placeholder for any *variate* (in contrast to constituent of natural reality) constructed so as to satisfy the requirements of common factor-hood as specified by the equations of linear factor analysis. Neither the cpc, nor its properties, are variates. What would it mean if the scalp recording data were unidimensional and the scores on one of the common factor variates turned out to have a particular pattern such as (1,-1,1,-1,...)? It would mean that it exhibits an agreement with the sequence of pulses that were the cause of the electrical activity produced by the cpc. With respect our understanding of the CPC, this is hardly an important fact.

If we merely *believed* that the electrical activity recorded on the scalp emanates from a heretofore unknown brain structure, then neurological, not factor analytic, investigation is the route to the discovery of the brain structure. And, if we did discover an existing material cpc, then any of a range of detectors of electrical fields currently in existence could be used to detect its electrical output on the surface of the scalp. The dimensionality of the scalp measurements as told by a factor analysis would have no direct implications with respect the issue of the existence or detection of such an hypothesized material structure. If, based on theory and knowledge, various possible laws (mathematical equations) were deduced about the relationship between an hypothesized to exist cpc and the scalp recordings, then these might well bare on the question of existence of a cpc (just as theory from physics implied that neutrinos should exist). However, the equations of factor analysis are not such possible laws. For they do not relate, say, the voltage produced by the cpc, to the dimensionality of the scalp recordings, but, rather, properties of a *random variate* (the latent variate) to the dimensionality of the scalp recordings.

It might be argued, however, that while a linear factor analysis of the scalp recordings can't *settle* the issue of the existence of an hypothesized cpc, a unidimensional result must at least give researchers *confidence* that such a cpc exists, and *inspire* them to search for it in the brain. It will be remembered that some of Mulaik's recommendations are very much in this spirit. But is there anything to this? If the dimensionality of the scalp recordings turned out to be undimensional, then real, ordinary empirical investigations into the brain would be required to prove the existence, and, if it exists, study the properties, of such a cpc that is currently believed to exist. And for this investigation to bear fruit, science would have to antecedently lay down rules of correct employment of at least a rudimentary concept *cpc*. Without clarity in the grounds of application of this concept it would have no way to *identify* some constituent *x* of the brain *as* the previously hypothesized and now discovered cpc. If, on the other hand, the dimensionality of the scalp recordings was not unity, but science believes that a cpc exists, then it must nonetheless push ahead with real, ordinary empirical investigations into the brain in the hope that a cpc can be found. Multidimensionality of the scalp recordings can, after all, arise for all kinds of reasons. If a cpc were eventually to be discovered, then research into its properties would be *required* to explain the original multidimensional linear factor analytic result. The dimensionality of the scalp recordings can have no bearing on an issue of existence of an hypothesized brain structure. Thus, regardless of what a linear factor analysis says about the dimensionality of the scalp recordings, when science believes that a cpc exists, real, ordinary empirical investigations into the brain will be the only arbiter of the truth of this belief. What exactly did the factor analysis add?

The misportrayal of factor analytic results as bearing on the existence of a cpc is *not* an example of affirming the consequent, because the fundamental material implication of linear factor analysis does not take as the antecedent, "the electrical activity is produced by the cpc", and as the consequent, "the scalp electrical measurements are unidimensional in a linear factor analytic sense." If this *were* the material implication in play, and it were true, then it would be called a law. Note that such laws can be generated only given the usual requirements: i) the capacity to identify in nature the relata of the law (a conceptual capacity); ii) intensive study of the relata, this presupposing the existence of each relata, yielding knowledge of their properties, locations, and relationships. If this material implication were known to be true, then, of course, it would have nothing to do with latent variable modelling: the antecedent would not represent a latent variate, but, rather, an existing brain structure known as the cpc. That is, the establishment of a law relating the electrical output of a brain centre and the dimensionality of scalp electrical recordings is the product of research, not latent variable modelling. The material implication that is fundamental to linear factor analysis takes as the antecedent, "there exists a continuous random variate $\boldsymbol{\theta}$ with $E(\boldsymbol{\theta})=0$, $V(\boldsymbol{\theta})=E\boldsymbol{\theta}^2=1$, and such that the p residuals, $\underline{\mathbf{l}}$, of the linear regressions of the input variates, $\mathbf{X}_j$, j=1..p, on $\boldsymbol{\theta}$ have a covariance matrix that is diagonal and positive definite", and the fact that this antecedent specifies properties that a *random variate* must have in order to be justifiably called a common factor reveals the fact that latent variable models are not detectors, but, rather, random variate generators.

Scenario 2:

Let's say that by the phrase "belief that there is a color centre" we instead mean "belief that there is an unobservable (i.e., *in principle* unobservable) cpc." Then we have here left the domain of science and stepped into the metaphysics of the Central Account. Much of the rest of the scenario would then become "mere words", the use of science talk to mask a metaphysics. For example, given that the cpc is *in principle unobservable*, then it is not a material structure. No material structure is "in principle" unobservable. Material structures do not reside within "latent domains." Material entities that are truly unobservable are unobservable for a reason, and scientists spend a great deal of time trying to find ways to render them observable. And because the cpc is not truly a material structure, it has no location within the brain (the constituents of the brain are of a material nature). Hence, there can be no true technical problem pertaining to our "inability to insert an electrode into the cpc." Because metaphysically unobservable entities do not generate electrical signals (only constituents of natural reality do), the cpc is not something whose electrical activity *could* be detected. This is why we are, under this scenario, attempting to portray a little set of equations (those of linear factor analysis) as a detector of this mysterious metaphysical cpc. The electrical outputs of things that exist and produce electrical outputs are detected by tools constructed to detect electrical fields, and sets of equations do not number among these.

If the cpc is unobservable in principle, and, hence, a metaphysical entity, what can a factor analysis say about such a thing? The answer is obviously, "nothing", because there is no such thing as detecting a metaphysically unobservable thing. What if the data from the p scalp electrodes turned out to be unidimensional in a linear factor analytic sense. Well, this would be an empirical fact about a set of variates. It would mean that they were replaceable in a sense of replaceability unique to linear factor analysis, and described in Part III of this book. Factor analysis can say nothing more. If, in fact, the electrical activity recorded on the scalp did emanate from a single material brain structure, then we must return to the science of scenario one to find this out.