PREDICTION IN SOCIAL SCIENCE
The Case of Research on the Human Resource Management-Organisational Performance Link[1]

BY

STEVE FLEETWOOD & ANTHONY HESKETH

Abstract. Despite inroads made by critical realism against the 'scientific method' in social science, the latter remains strong in subject-areas like human resource management. One argument for the alleged superiority of the scientific method (i.e. its scientificity) lies in the taken-for-granted belief that it alone can formulate empirically testable predictions. Many of those who employ the scientific method are, however, confused about the way they understand and practice prediction. This paper takes as a case study empirical research on the alleged empirical association between human resource management practices and organisational performance. By unpacking the confusion surrounding the two basic notions of prediction used, it reveals what is wrong with them, why the scientific method cannot actually make accurate predictions and why, therefore, the scientific method fails to meet its own criteria for scientificity. Finally, explanation is considered in order to prevent any confusion between it and prediction and to offer what we call *tendential prediction*.

Key words: explanation, human resource management, performance, prediction, scientism, tendential prediction

Introduction

Despite inroads made by critical realism against the use of what is often referred to as 'the scientific method' in social science, the latter remains remarkably strong in subjects like human resource management (HRM), economics and psychology. One of the most powerful arguments for the alleged superiority of the scientific method over other methods (i.e. its scientificity) lies in the taken-for-granted belief that *it alone can formulate empirically testable predictions*. Many of those who employ the scientific method (along with many who simply opt for other methods) are, however, extremely

[1] We would like to thank two anonymous referees for comments that helped us rework, and improve, the paper.

© Koninklijke Brill NV, Leiden, 2006

Also available online – www.brill.nl

equinoxonline

confused about the way prediction is understood and practiced. This makes it harder for critics to grasp the understandings and practices involved and, therefore, harder to formulate a critique. As a result, critics are often charged with inventing straw persons. This paper takes empirical research on the alleged link between HRM practices and organisational performance as a case study of the scientific method in action. It reveals (some of) the confusion with which prediction is understood and practiced in the field. The lessons drawn are, however, applicable to almost any branch of social science where aspects of the scientific method are used.

The paper has four parts. Part 1 provides a thumbnail sketch of the HRM-P paradigm and clarifies some issues surrounding the notions of science (or scientism) and prediction. Part 2 considers the meta-theory underpinning research on the HRM-P link and uncovers exactly why the HRM-P system is an open system. Part 3 unpacks the confusion surrounding the two notions of prediction that lie buried in the literature. Each of these notions is then scrutinised to reveal what is wrong with them, why the scientific method cannot actually harness the power of prediction and why, therefore, the scientific method *fails to meet its own criteria for scientificity*. The concluding part briefly considers explanation to prevent any confusion between it and prediction.

1. *Some Preliminaries*

The HRM-P paradigm

HR professionals are desperately trying to demonstrate that HRM adds, rather than saps, value, and have turned to research from consulting houses[2]

[2] For example: Watson/Wyatt, *Human Capital Index®: Human Capital As A Lead Indicator of Shareholder Value*, pp. 1-5, 2001/2; R. Donkin, 'The rise of people power', *Financial Times* 29 April, 2004; S. Overell, 'The metric system of performance', *Financial Times* 31 October, 2002; Task Force on Human Capital Management, *Accounting for People*, Office of the Secretary of the State for Trade and Industry, 2003; R. Thomas, P. Cheese and J. Benton, 'Human capital development', *Accenture Research Note* 1, 2003, pp. 1-4; PricewaterhouseCoopers, *Sustaining Value Through People*, PriceWaterhouseCoopers publication, 2003, p. 9; W. Bukowitz, R. Williams and E. Mactas, 'Human capital measurement', *Industrial Research Institute* May-June, 2004, pp. 43-49; Accenture, 'The High Performance Workforce Study', Accenture publication, 2004, pp. 1-78; M. Witzell, 'A better feel for the intangibles', *Financial Times* 29 April, 2004; Saratoga, *Key Trends in Human Capital: A Global Perspective*, PriceWaterhouseCoopers publication, 2005, pp. 1-19;

equinoxonline

and academics. The over-riding message emanating from this voluminous research is that a measurable, empirical association exists between an organisations' HRM practices and its performance—henceforth referred to as the HRM-P link. As we will see, prediction plays a key role here.

From science to scientism

Supporting and sustaining research on the HRM-P link is what is often referred to as a 'scientific' approach. Boudreau and Ramstad refer to 'scientific studies';[3] Murphy and Zandvakili suggest that 'scientific measures be used to evaluate the effectiveness of HRM practices' referring to 'data collected by scientific methodology';[4] Brown refers to the 'science of human capital measurement'[5] and Thomas and Burgman suggest that human capital management is moving 'from art to science'.[6]

We are not aware of any researchers in the HRM-P paradigm who have reflected upon, or defined, the scientific method they use exclusively. This is likely to be because when most postgraduate researchers learn 'methodology', they are usually presented with a set of statistical techniques (usually with little or no discussion of serious methodology and/or philosophy of science), as if these techniques just are the only ones available for anyone who wants to be 'scientific'. This taken-for-granted attitude is illustrated by Booth's[7] work on the economics of trade unions, where she refers to 'the accepted methodology of economics', without feeling the need to actually state what it is.[8]

People, Performance, Profit, Montgomery Research and Accenture, 2005; *Human Capital Management*, People Management, in association with Oracle, 2005.

[3] J. Boudreau and P. Ramstad, 'Measuring intellectual capital: learning from financial history', *Human Resources Management*, vol. 36, no. 3, 1997, pp. 343-56, p. 343.

[4] T. Murphy and S. Zandvakili, 'Data and metrics driven approach to human resource practices: using customers, employees and financial metrics', *Human Resource Management*, vol. 39, no. 1, 2000, pp. 91-105.

[5] D. Brown, 'Capital vetters', *People Management*, 30 September, 2004, pp. 38-41, p. 40.

[6] R. Thomas and R. Burgman, 'Managing human capital as an asset', *Accenture Research Note*, vol. 1, 2005, pp. 1-4.

[7] A. Booth, *The Economics of Trade Unions*, Cambridge: Cambridge University Press, 1995, p. 83.

[8] In a recent survey of 467 articles on HRM, Hoobler and Brown Johnston, not only found just *one* article on meta-theory, they also found that 'statistical regression was by far the method of choice, represented in a full 35 percent of the articles studied. Various analysis of variance and meta-analysis accounted for 9 percent and 5 percent

equinoxonline

To give a flavour of what this accepted scientific method entails in the HRM-P paradigm, we sketch it as follows: *Although variations exist in the phenomena that are measured, and the metrics and measures that are used to quantify these phenomena, HRM practices and organisational performance are quantified and empirical data generated. Various statistical techniques (typically regression, analysis of variance, correlation, structural equation modelling and factor analysis) are then employed on this quantitative data to empirically test various predictions (or hypotheses) to the effect that certain bundles of HRM practices lead to increased organisational performance.*

Critics like us, however, argue that 'scientism' (or derivates such as 'scientistic') is a more appropriate definition of the method used in research on the HRM-P link.[9] The *Collins Dictionary of Sociology* defines scientism as 'any doctrine or approach held to involve oversimplified conceptions and unreal expectations of science, and to misapply "natural science" methods to the social sciences'.[10] Hughes and Sharrock define scientism as 'those philosophies such as positivism, which seek to present themselves as having a close affiliation with the sciences and to speak in their name, and which then go on to fetishise the so called scientific standpoint'.[11] For us, then, a perspective is scientistic if it loosely refers to the employment of methods and techniques allegedly similar to (some aspects of) natural science, without actually specifying what these methods and techniques are and why they are appropriate to social science.[12]

Centrality of prediction

Central to scientism is prediction—although this is due less to careful reflection, and more to a kind of taken-for-granted belief that the science is primarily about formulating and testing predictions and not necessarily

---

respectively, while correlation and structural equation modeling or confirmatory factor analysis respectively amounted to 6 percent and 3 percent'. (J. Hoobler and N. Brown Johnston, 'An analysis of current HRM publications', *Personnel Review*, vol. 33, no. 6, 2005, pp. 665-6, p. 668).

[9] Indeed, we prefer the term 'scientism' to 'positivism' on the grounds that (a) positivism comes in many shapes and sizes and we want to avoid getting sidetracked by definitional issues, and (b) many genuine positivists are perfectly aware of their method as they have reflected carefully upon it.

[10] *The Collins Dictionary of Sociology*, Glasgow: Harper Collins, 1995.

[11] J. Hughes and W. Sharrock, *The Philosophy of Social Science Research*, London: Longman, 1997, p. 208.

[12] The argument in favour of defining research on the HRM-P link as 'scientistic' will become stronger after part three when its various claims are exemplified.

equinoxonline

about meeting objectives like realisticness. Friedman's seminal paper on prediction as the sole criteria for evaluating theories has been enormously influential in grounding this belief.[13] We can think of four main reasons for this centrality.

(1) One of the most powerful arguments for the alleged superiority of the scientific method over other methods (i.e. its scientificity) lies in the taken-for-granted belief that *it alone is able to formulate empirically testable predictions*. Perspectives that do not, or cannot, do this and aim instead for things like the recovery of actors' meaning (interpretivists, hermeneuticists, ethnomethodologists), the deconstruction of phenomena as texts (postmodernists or poststructuralists), the analysis of discourse (critical discourse analysts) and/or explanation of the causal mechanisms that actors interact with (critical realists[14]), are presumed to be un-scientific and, in this sense, inferior.
(2) Methods that generate theories whose predictions can be empirically tested raise the possibility that some of these predictions will be successfully tested, thereby providing a basis for policy prescription. If we can predict the future outcome of an action, we may be able to initiate that action, indeed, prevent that action, in order to bring about desired outcomes.
(3) In some natural sciences (typically those whereby the system under investigation is spontaneously closed or can be closed easily), successful predictions can be made. This success encourages the belief that, if social scientists continue to follow the example of these 'mature' sciences, one day, social sciences like HRM too will be able to make successful predictions. In the meantime, we should continue our efforts to generate successful predictions.
(4) Drawing on the work of Tsang and Kwan, we might say that prediction is superior to its close relative *accommodation*.[15] A researcher who

[13] M. Friedman, 'The methodology of positive economics', in D. Hausman, *The Philosophy of Economics: An Anthology*, Cambridge: Cambridge University Press, 1988.

[14] It is instructive to note that Chadwick and Capelli actually mention 'causal mechanisms' but then proceed immediately to a discussion of how to measure them. (C. Chadwick and P. Capelli, 'Alternatives to generic strategy typologies in strategic human resource management', *Research in Personnel and Human Resource Management* supplement 4, 1999, pp. 1-29).

[15] E. Tsang and K.-M. Kwan, 'Replication and theory development in organisational science: a critical realist perspective', *Academy of Management Review*, vol. 24, no. 4, 1999, pp. 759-80.

equinoxonline

> constructs theory to fit the data, accommodates the data. Another researcher may use this theory to make and test a prediction. Accommodation can be fudged, that is, the researcher knows the result the theory should generate and fudges the theory to make the theory fit the data. In the case of prediction, however, the theory comes into existence before the data and cannot be fudged.

Having established that research on the HRM-P link is usefully described as scientistic, having a grasp of what it entails, and recognising why prediction appears central, we can shift our focus towards understanding the meta-theory that underpins all this.

2. *Meta-Theoretical Underpinnings of Research on the HRM-P Link*

We intend here to be relatively brief in invoking the (now fairly 'standard') critique of closed systems, while also setting out reasons why the particular system we are interested in, the HRM-P system, is actually open.

Scientism's generally accepted method appears to be some (unspecified) variant, or combination, of the covering law model, deductive nomological model, inductive-statistical model, or hypothetico-deductive method. Following Lawson critical realists have referred to this variant as the *deductive method* or simply *deductivism*.[16] From this perspective to 'explain' something is to predict a claim about that something as a deduction from a set of initial conditions, assumptions, axioms, and law(s) or some other regular pattern of events.

Scientism presupposes (explicitly or implicitly) an ontology consisting of what can be observed and is, therefore, of observed *events*. Because these objects are confined to experience the ontology is *empirical*; and because these objects are thought to exist independently of one's identification of them, it is *realist*. The ontology is, therefore, *empirical realist*. If *particular* knowledge is gained through observing events, then *general*, including *scientific*, knowledge is only available if these events manifest themselves in some kind of pattern: a flux of totally arbitrary events would not result in knowledge. Scientific knowledge is, therefore, entirely reliant upon the existence and ubiquity of event regularities or constant conjunctions of events—we use these phrases interchangeably.

[16] T. Lawson, *Economics and Reality*, London: Routledge, 1997.

equinoxonline

Critical realists typically generalise and style regularities between events as 'whenever event *x* then event *y*' or 'whenever event $x_1 \ldots x_n$ then event *y*'. Regularities between variables are more often expressed as functional relations, $y = f(x)$ or $y = f(x_1 \ldots x_n)$ and this is the way they appear in, for example, regression models.

Research on the HRM-P link is preoccupied with what is referred to variously as testing the prediction, testing the hypothesis, testing the theory, testing the model, testing the model's predictions, finding the predictors of their dependent variable and so on. The terminology varies and, it must be said, is highly ambiguous, but the practice is well known. In what follows we will refer (where possible) to *testing the hypotheses*. A hypothesis is a very precise statement about what will regularly happen to the *magnitude* of one variable when the *magnitude* of another variable or variables occurs or changes. The key points, however, are that predictions and hypotheses are (a) only intelligible if they are expressed in terms of *regularities* between events or variables; and (b) only possible if event regularities are ubiquitous. Predictions and hypotheses are only intelligible and possible, if event regularities exist, and event regularities occur in closed systems. Let us consider closed and open systems in more depth.

Closed and open systems

Whilst there are several ways to define systems, critical realists define systems as closed when they are characterised by event regularities, and open when characterised by a lack of such regularity.[17] Events are constantly conjoined in the sense that for every event y, there exists a set of events $x_1, x_2 \ldots x_n$, such that y and $x_1, x_2 \ldots x_n$ are regularly conjoined. A deterministically closed system can be expressed probabilistically and can, thereby, be transposed to a stochastically closed system. Here y and $x_1, x_2 \ldots x_n$ are regularly conjoined under some well-behaved probabilistic function. In effect, the claim ' whenever event x then event y' is transposed into the claim 'whenever events $x_1, x_2, \ldots x_n$ on average, then event y on average', or 'whenever the average value of events measured by variables $x_1, x_2, \ldots x_n$ are what they are, then the average value of event y measured by variable y is what it is'. Stochastically closed systems, are still closed systems.

The important point to note here is that without event regularities, that is to say, in open systems, prediction based upon inductive generalisation

[17] On various definitions of open and closed systems, see the exchange between Fleetwood and Dow in *Journal of Critical Realism*, vol. 5, no. 1, 2006.

equinoxonline

is not possible. If it is *not* the case that event y is observed to regularly follow events $x_1, x_2, \ldots x_n$ then we have no grounds for the inductively generated prediction that the next time events $x_1, x_2, \ldots x_n$ occur, event y will follow.

Does the HRM-P literature presuppose a closed system? In a word: yes. To suggest, as the literature overwhelmingly tries to, that some HRM practices are statistically associated with increased performance, is to assume regularity and hence closure. If textual evidence is needed, the following influential commentator even uses terminology that could be lifted straight from virtually any critical realist account of closed systems:

> Ideally, you will develop a measurement system that lets you answer questions such as, how much will we have to change x in order to achieve our target in y? To illustrate, if you increase training by 20 percent, how much will that change employee performance and, ultimately, unit performance?[18]

Whilst constant conjunctions of events and, therefore, closed systems, are fundamental to deductivism, they are exceptionally rare phenomena. There appear to be very few spontaneously occurring systems wherein constant conjunctions of events occur in the natural world, and virtually none in the social world. This is not to deny the possibility that constant conjunctions may occur accidentally, or over some restricted spatio-temporal region, or be trivial. But virtually all of the constant conjunctions of interest to science only occur in artificially closed systems, typified by the bench experiments of some natural sciences. In those natural sciences where experiments can be carried out, the point of the experiment is to close the system by engineering a particular set of conditions that will isolate the one interesting mechanism. This mechanism is then allowed to operate unimpeded and the results, the constant conjunctions, recorded. In social science, however, constant conjunctions only occur where they are engineered in the form of theoretically closed systems.

This throws up two questions. *Why are social systems open*? and *What are the consequences of modeling open systems as if they are closed*? It is difficult to present these two questions sequentially, so each one will be discussed when it is appropriate to do so.

Let us start by considering the social system of interest to us—the workplace—and why it is probably an open system. The workplace can be thought of as a system where HRM practices, in part, cause changes in

[18] B. Becker, M. Huselid and D. Ulrich, *The HR Scorecard: Linking People, Strategy and Performance*, Harvard: Harvard Business School Press, 2001, p. 110.

equinoxonline

organisational performance. If it transpires that whenever some bundle of HRM practices are introduced into the workplace, organisational performance changes (let us say increases), then we would be dealing with a closed system. If no such regularity occurred, then we would be dealing with an open system. It will come as no surprise to anyone with some familiarity of social systems like workplaces that they are *multiply caused, complex, evolving* and subject to the *exercise of human agency*. Moreover, we strongly suspect our claim that the social world is characterised by multiple causality, complexity, evolution and human agency is not controversial and would be accepted even by researchers in the HRM-P paradigm. Let us consider the workplace in a little more detail.

- The system is multiply caused, in the sense that there are probably scores or even hundreds of ensemble (and sub-practices and sub, sub-practices etc.) that have some kind of causal impact on organisational performance. Their causal impact can be *direct* in the sense that each individual practice directly effects performance; and/or *indirect* in the sense that each ensemble interacts with other practices in the bundle, the bundle generates synergy, and this synergistic bundle then effects performance. Recognising multicausality requires more than simply trying to disaggregate the variables or trying to add in as many variables as it is possible to obtain data on, or trying to find 'missing variables'. This is, in part, because many of the causal factors may be unobservable, or if observable, impossible to meaningfully quantify and reduce to variables, so are simply left out (literally) of the equation. It is also, in part, because of complexity and evolution.
- The system is complex in the sense that it generates its own internal changes which feedback to alter the nature of the HRM practices, and this changes the effects these practices (directly and indirectly) have on performance. Complexity introduces difficulties in reducing something to a variable if that something is undergoing a change in its nature.
- The system evolves in the sense that it is always creating and responding to changes in the external environment and, once again, these changes alter the nature of the HPW practices, and this changes the effects these HPW practices (directly and indirectly) have on performance. This often leads to causal effects being intermittent in the sense that HPW practices operating a certain way today, may 'switch off' as it were, or start to operate in slightly different ways, tomorrow only to 'switch on' again or revert back to their old mode of operating at some later date. Evolution

equinoxonline

also introduces difficulties in reducing something to a variable if that something is undergoing a change in its nature—we will return to this below.

- The system is subject to human agency in the sense that human beings can, and do, change their minds. This should not be taken to mean that humans are entirely capricious or act whimsically. Rather, it means their actions are not entirely predictable and they retain the ability to always have done otherwise. To deny this is to deny human subjectivity, creativity, imagination, ingenuity and entrepreneurial activity. HPW practices that were accepted as legitimate by workers in one period can become unacceptable in another period and vice versa, and it is often difficult to attribute causes to this other than to say workers changed their minds.
- To claim that workplaces are multiply caused, complex, evolving and subject to the exercise of human agency, is not to claim that these systems continually undergo total transformation so that in each period we are faced with a kind of radically new system. Systems, like workplaces, are usually characterised by a form of quasi-reproduction (reproduction with slight variation) and occasionally by radical transformation. For example, whilst the set of implicit rules that govern industrial relations arrangements are (quasi) reproduced, so that today's arrangements resemble yesterday's, continual conflicts and tensions may generate slight variations in the rules, leading to a lack of (or increase in) trust between managers and union lay-officials. Occasionally, of course, the set of implicit industrial relations are radically transformed by things like union de-recognition or mass redundancies.

Whilst we do not have to uncritically swallow the motifs of management discourse—wherein 'the only constant is change', and we are all 'surfing on chaos' or whatever—many of these management commentators make ontological presuppositions that are far more in tune with the way the world is than those of a scientistic persuasion. Brown and Eisenhardt make plausible claims about much contemporary business being: *unpredictable and uncontrolled*. 'The future is too uncertain for such pin-point accuracy [. . . and . . .] there is simply too much going on in rapidly changing industries for a single group to orchestrate every move'.[19] A recent study from *PriceWaterhouseCoopers* notes that 'the relationship between historic trends

[19] S. Brown and K. Eisenhardt, *Competing On The Edge: Strategy as Structured Chaos*, Harvard: Harvard Business School Press, 1998, p. 8.

equinoxonline

and future performance is not always well understood, in part, because people (all stakeholders, not just employees) behave in different, sometimes unpredictable ways [. . .] therefore, the diverse effects of people policies on business value may be both complex and, in some cases unexpected'.[20]

If we have grounds to believe that the workplace is characterised by multiple determination, complexity, evolution, and is subject to the exercise of human agency, then we have strong a priori grounds to believe that event regularities are most unlikely to occur, and the system is open. We would, for example, expect to find that when bundles of HRM practices are introduced into workplaces, sometimes actual performance improves (a little or a lot), sometimes it remains unchanged and sometimes it deteriorates (a little or a lot). Moreover, we would not expect to be able to predict which of these outcomes will prevail. What does the empirical evidence show?

Rather than trawl through the literature, which is extensive, we cite three recent surveys of research seeking an empirical association between HRM and performance. Wall and Wood[21] conclude that 'existing evidence for a relationship between HRM and performance should be treated with caution.' Godard writes: 'Overall, these concerns suggest that we should treat broad-brush claims about the performance effects of [High Performance Work systems], and about research findings claiming to observe them, with a healthy degree of scepticism.'[22] Boselie, Dietz and Boon conclude that:

> A steady body of empirical evidence has been accumulated since the pioneering days of the mid-1990s, yet it remains the case that no consistent picture exists on what HRM is, or what it is supposed to do [. . .] What can be concluded definitively from this collection of studies is still unclear [. . .] Ten years on the 'Holy Grail' of decisive proof remains elusive.[23]

---

[20] PriceWaterhouseCoopers, *Sustaining Value Through People*, PriceWaterhouseCoopers publication, 2003, p. 9. This study seems not to spot the contradiction it highlights. There simply may not be a 'relationship between historic trends and future performance' in part 'because people [. . .] behave in different, sometimes unpredictable ways. Hence, the idea that this alleged relationship is 'not always well understood' (p. 26) misses the point completely. The report confuses the (reasonable) desire of firms to be able to predict the future with their ability to do so.

[21] T. Wall and S. Wood, 'The romance of human resource management and business performance, and the case for big science', *Human Relations*, vol. 58, no. 4, 2005, pp. 429-62.

[22] J. Godard, 'A critical assessment of the high-performance paradigm', *British Journal of Industrial Relations*, vol. 42, no. 2, 2004, pp. 349-378, p. 355.

[23] P. Boselie, G. Dietz and C. Boon, 'Commonalities and contradictions in HRM

equinoxonline

In sum, we not only have strong a priori grounds to support the contention that the workplace is an open, not a closed system, the empirical evidence suggests this conclusion.[24] Let us now turn to consider how systemic openness robs deductivism of predictive power.

3. *Can Scientism Formulate Empirically Testable Predictions?*

Some researchers are prepared to state quite boldly that their findings have predictive uses:

> The results indicate that [. . .] the direct effect on the sales volume is $19.30 for each dollar invested in customer services, and $25.50 for each dollar spent in sales training . . . *The approach may also be used for forecasting* future training ROI [returns on investment].[25]

Wang, Dou and Li do not, however, do this or state how it can be done. Ahmad and Schroeder claim that canonical correlation 'can be used for predictive purposes'[26]—ignoring the fact that statistical association such as correlation does not imply causality and without the later it seems difficult to know how predictions could be made. For Becker, Huselid and Ulrich measurement is closely connected to prediction because:

> it improves HR decision making by helping you focus on those aspects of the organisation that create value. In the process, it provides you with feedback that you can use [. . .] to predict the impact of future decisions.[27]

Huselid, whose 1997 paper is regarded by Godard[28] as 'the most noted study', appears to use his measures to facilitate prediction:

---

and performance research', *Human Resource Management Journal*, vol. 15, no. 3, 2005, pp. 67-94, pp. 81-2.

[24] In coming to this conclusion, we have taken the 'usual' statistical procedures used in empirical research on the HRM-P link without criticism, ignoring the fact that we are highly critical of many of these procedures. All we have established is that, even taking these procedures at face value, we have good reason to believe that the workplace is more appropriately characterised as an open, not a closed system.

[25] G. Wang, X. Dou and N. Li, 'A systems approach to measuring return on investment for HRD interventions', *Human Resource Development Quarterly*, vol. 13, no. 2, 2003, pp. 203-224, p. 217, emphasis added.

[26] S. Ahmad and R. Schroeder, 'The impact of human resource management practices on operational performance: recognising country and industry differences', *Journal of Operations Management*, vol. 21, 2003, pp. 19-43, pp. 27-8.

[27] Becker, Huselid and Ulrich, *HR Scorecard*, p. 110.

[28] J. Godard, 2004, 'A critical assessment', p. 352.

equinoxonline

> To estimate the practical significance of the impact of High Performance Work Practices on productivity, I next calculated the impact of a one-standard deviation increase in each practice scale on [. . .] net sales [. . .] The findings indicate that each one-standard-deviation increase *raises* sales an average of $27,044 per employee.[29]
>
> A one-standard deviation increase in such practices is associated with a relative 7.05 percent decrease in turnover and [. . .] $27,044 more in sales and $18, 641 and $3,813 more in market value and profits respectively.[30]
>
> These [. . .] values suggest that firms *can indeed obtain* substantial financial *benefits* from investing in the practices studied here.[31]
>
> This study also provides one of the first *tests of the prediction* that the impact of High Performance Work Practices.[32]

Careful reading (of this and other studies in the literature), reveals that *two* notions of prediction are buried within research and, unfortunately, they are not always carefully disentangled. The following section tries to untangle them.

Untangling two notions of prediction

The first, and most *discursively powerful*, notion uses prediction in the sense of *predicting a future* event or state of affairs. Most empirical research, such as that on the HRM-P link is littered with phraseology like 'predicting the impact of future decisions'; 'practical significance'; 'an increase in HRM raises sales'; and 'firms can indeed obtain benefits' suggest the use of prediction in this sense. Let us refer to this as $prediction_f$—with the subscript denoting 'future' or possibly 'forecast'. Note well that $prediction_f$ is what really carries the weight of scientism's alleged superiority, appears to deliver the killer blow to alternatives; this is why we refer to it as discursively powerful. To be sure, if a theory predicts that X will occur tomorrow, and X does occur tomorrow, this theory will be considered to be a good example of the scientific (or scientistic) method.

The second, and less discursively powerful, notion uses prediction in the sense of *testing hypotheses* via events or states of affairs that have already

[29] M. Huselid, 'The impact of human resource management on turnover, productivity and corporate financial performance', *Academy of Management Journal*, vol. 38, no. 3, 1995, pp. 635-672, p. 658, emphasis added.

[30] Ibid., p. 667.

[31] Ibid., emphasis added.

[32] Ibid., p. 636, emphasis added.

equinoxonline

occurred and are now in the past. This is variously described as ascertaining data consistency or fitting a model. From past data, on past phenomena, we deduce or 'predict' a past outcome. While this outcome has already occurred it could have been predicted from the data had we done so at an earlier time. Let us refer to this as *prediction*$_p$—with the subscript denoting 'past'. Prediction$_p$ might be less discursively powerful than prediction$_f$, but it still carries connotations of being scientistic and, therefore, retains a degree of discursive power.

Now, the problem is that research on the HRM-P link never quite specifies which notion of prediction is being used at any moment. Huselid's work, for example, displays elements of both.

Consider the case for prediction$_p$. Amongst others, Huselid tests the hypothesis that High Performance Work Practices (HPW) will increase sales and profits. He obtains data by recording *past* instances where HPWs were in use and *past* instances where sales and profits changed. If a (significant and positive) statistical association is found in the data between these events of the past, he suggests that the data confirms (or does not falsify) the hypothesis. This is sometimes referred to as *data consistency*: the data are consistent with events from which they are drawn.

Consider the case for prediction$_f$. Huselid does not go as far as saying: 'my findings allow me to predict$_f$ that if your firm introduces these practices, then your firm will enjoy decreases in turnover and increases in sales and profits of something like these magnitudes.' Yet something like prediction$_f$ is not only implied, it follows quite naturally from the hypotheses his research tests. If prediction$_f$ is not implied, then his findings have no practical significance, and one of the key features of his paper is lost.

Huselid uses the more discursively powerful notion of prediction$_f$ whilst actually *practicing* the less discursively powerful notion of prediction$_p$. As far as we are aware, Huselid has not actually made any kind of predictions$_f$, despite phraseology to the contrary. But, then again, neither has anyone else in the literature. *Our point here is not to criticise Huselid (and others) for not engaging in prediction*$_f$*, but to illuminate the ambiguity surrounding the term 'prediction'.* This ambiguity allows advocates of the scientistic perspective to harness the discursive power of prediction$_f$ whilst actually practicing prediction$_p$ which is less discursively powerful. In this way, the scientistic credentials of deductivism are enhanced.

In sum, whilst research on the HRM-P link often appears to make predictions$_f$, in reality it almost always makes predictions$_p$. But surely, if this less discursively powerful notion still harnesses the power of prediction, can

equinoxonline

it not be said to be superior to alternatives? To answer this question, we need to unpack the notion of prediction_p.

Imagine we conducted a typical piece of research and predicted_p that the existence of teamwork and incentive pay will be associated with increased productivity. Suppose this prediction_p turns out to be data consistent, allowing us to say teamwork and incentive pay are good predictors_p of increased productivity. What we have is a consistent prediction_p from a set of data about a firm or a sample of firms. And this is, typically, where most empirical research (on the HRM-P link or otherwise) ends. Now, imagine we conducted a typical piece of research on cold fusion. Imagine we predicted_p that the existence of substance S_1 and substance S_2 (in a test tube) are associated with cold fusion. Suppose this prediction_p turns out to be data consistent. What we have is a consistent prediction_p from a set of data about a single experiment. Unlike the previous case, however, research would not end here. Other researchers, from within this scientific community would seek to *replicate* these findings—indeed this is exactly what happened in the case of alleged cold fusion a few years ago. And if in another experiment(s) the prediction_p turns out to be data consistent, then the theory that led us to combine substances S_1 and S_2 is confirmed, or not-falsified.

The point of this imaginary exercise is to introduce the notion of *replication*. Until and unless prediction_p is replicated all we have is the equivalent of a single experiment, and this would not be accepted in the kinds of natural sciences for which experiments are possible, and from which scientism in the social sciences draws strength. Does research on the HRM-P link engage in replication? The answer to this question is not only 'No it *does* not', but also, 'No it *can*not'.

Replication

Replication is not a straightforward notion. Tsang and Kwan[33] identify six types of replication:

- *Checking of analysis*. Subsequent researchers employ exactly the same measurement, analysis and data set
- *Reanalysis of the data*. Subsequent researchers employ different measurement and analysis but exactly the same data set.

[33] Tsang and Kwan, 'Replication'.

equinoxonline

- *Exact replication*. Subsequent researchers employ exactly the same measurement and analysis, on the same population, but a different sample and hence different data set. This is done to assess whether the findings are reproducible.
- *Conceptual extension*. Subsequent researchers employ different measurement and analysis on the same population, but a different sample and hence different data set. Subsequent models are extended to include different causal mechanisms, or different variables.
- *Empirical generalisation*. Subsequent researchers employ exactly the same measurement and analysis on a different population, different sample and hence different data set. This is done to assess whether the findings are generalisable to another population.
- *Generalisation and extension*. Subsequent researchers employ different measurement and analysis but on a different population, different sample and hence different data set.

It should be noted here that Tsang and Kwan are discussing replication in epistemological, rather than ontological, terms. That is, although subsequent researchers might employ different measurement, analysis, population, sample and data set, there is no suggestion that the phenomena under investigation are different. They are discussing different ways of gaining knowledge (epistemic) of the same relatively unchanging phenomena (ontic). This will become important below.

Understanding why replication does not, and cannot, be carried out in research on the HPWS-P link is difficult for the obvious reason that *we cannot point to examples of something that is not done!* We will proceed, therefore, by considering a hypothetical scenario of what would occur if one researcher ever attempted to replicate the findings of another researcher—or indeed attempted to replicate her own findings.

Let us assume, for argument's sake, that there is a theory explaining why three HRM practices (teamwork, performance related pay and flexible working practices) causes an increase in productivity. From theory T, Smith constructs the following model[34] with the HRM practices as independent (or 'explanatory') variables and productivity as the dependent variable:

$$Y = \alpha +, \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon \qquad (1)$$

[34] The point is easier to grasp if we do not use matrix notation.

equinoxonline

This is used to test the hypothesis$_1$ (or the prediction$_1$) that changes in teamwork, performance related pay and flexible working practices (X_1, X_2 and X_3 respectively) are associated with changes in productivity (Y). He collects data on the practices and productivity levels in a large UK organisation and estimates the coefficients. The usual diagnostic tests are run, the coefficients have the appropriate sign and are all significant. Smith concludes that the hypothesis$_1$ is not rejected by the data, and theory T gains support.

Two years later new data on HRM practices and productivity levels become available for the same large UK organisation that Smith investigated. Jones decides to use this new data to replicate the findings of Smith. This is what Tsang and Kwan referred to as 'exact replication.' She takes the same theory T that Smith used along with the same model and same hypothesis$_1$, and re-estimates the coefficients. It would, of course, be most unlikely if no changes had occurred in the organisation and its environment in the ensuing two years, it is not unreasonable to suppose there have been some slight changes. Smith finds that the previous model no longer fits the new data. Suppose X_2 is now insignificant and has the wrong sign. Jones faces two problems.

First, Jones has to deal with the thorny problem of what to do when theory T suggests the variable X_2 *should* be included, but the data suggests the variable *should not* be included. Does she stick with theory T and include a now insignificant and incorrectly signed variable X_2; or does she drop the variable? Dropping the variable implies there is something wrong with theory T. Smith is now in an awkward position because he has no theory to guide his choice of what else to measure and estimate. Incidentally, whilst this dilemma raises its head every day for empirical researchers (in the HRM-P paradigm and elsewhere) to the best of our knowledge, it is never recognised or addressed.

Second, let us suppose Jones proceeds by dropping X_2 from her new model. She then decides to include data on a new HRM practice, employee communication, (denoted X_4) that has been used within the organisation. She re-specifies the model thus:

$$Y = \alpha +, \beta_1 X_1 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon. \qquad (2)$$

Jones now re-estimates these coefficients. The usual diagnostic tests are run, the coefficients once again have the appropriate signs and are all significant. Jones concludes that the hypothesis$_2$ is not rejected by the data. It is now unclear what this says about theory. It is not a test of theory T.

equinoxonline

Moreover, because no theory guided the introduction of employee communication (X_4) and so there is no theory of why X_1, X_3 & X_4 might be related to Y, this is not a test of any theory whatsoever.

What has this hypothetical example established? Although Smith started out to perform an 'exact replication,' he ended up performing something more akin to 'generalisation and extension.' Tsang and Kwan,[35] question whether this is a replication or *a test of a different conceptual model*. Let us pursue this a little further.

Hypothesis$_2$ modelled by (2) is no longer the same hypothesis that was modelled by (1) because one of the variables contained in (2) was not present in (1) and one of the variables contained in (1) was not contained in (2). Whilst this may only be a small change, only affecting two variables, the number of variables involved here is irrelevant: the principle matters—we kept the example simple for ease of exposition. This is quite important and easily overlooked so allow us to put the point in other words. It is like saying:

- Hypothesis$_1$ and model (1) suggest that productivity increases are associated with teamwork, *performance related pay* and flexible working practices.
- Hypothesis$_2$ and model (2) suggest that productivity increases are associated with teamwork, flexible working practices and *employee communication*.

In order to test a hypothesis, both Smith and Jones had to specify the set of variables that constitute the model with absolute clarity: only one particular set of variables is associated with one particular hypothesis. If Jones changed one or more of the variables in the model, then she is, to be strictly accurate, specifying a *different* model.[36] *Jones would not have replicated Smith's initial study, she would, effectively have carried out another, different, study*.[37]

---

[35] Tsang and Kwan, 'Replication', p. 768.

[36] Ichniokski, Shaw and Prennushi measure HPW practices with 13 variables, Huselid measures them with 15 variables, and Ulrich suggests 69 possible measures. All that would be necessary is for one of these variables to change significantly for replication to be impossible. C. Ichniowski, K. Shaw and G. Prennushi, 'The effects of human resource management on productivity: a study of steel finishing lines', *The American Economic Review*, vol. 87, no. 3, 1997, pp. 291-313; Huselid, 'Impact'; D. Ulrich, *Human Resource Champions: The Next Agenda for Adding Value and Delivering Results*, Boston MA: Harvard Business School Press, 1997.

[37] Mirowski and Sklivas observe something similar in economics, claiming that most economists do not 'replicate' the findings of others, rather they 'reproduce' them.

equinoxonline

We noted above that if the workplace is characterised by multiple determination, complexity, evolution and is subject to the exercise of human agency, then it is *probably always undergoing changes*. This provides strong a priori grounds to believe that event regularities are most unlikely to occur within this system: the system is likely to be open, not closed. In open systems, prediction$_p$ cannot be used to test a hypothesis. To paraphrase Heraclitus, 'we cannot put our foot twice in the same system.' This takes us back to the point made about Tsang and Kwan discussing replication in epistemological, rather than ontological, terms. Replication in their taxonomy involves discussing different ways of gaining knowledge of the same unchanging phenomena. But we have seen that, in open systems, we would be most surprised to find the phenomena unchanged.

The problem for us, of course, is that we *do not* find cases of researchers attempting to replicate other studies, so we have to engage in this kind of hypothetical critique. Moreover, we are well aware that a hypothetical case does not prove that researchers on the HRM-P link *cannot* replicate one anothers' studies. If someone actually did succeed in replicating someone else's study, or even replicating their own study at a later date, then our point would be weakened. After all, the claim that the workplace is not characterised by event regularities, and is an open system, is ultimately an empirical claim even if we have strong a priori grounds to believe it.

A crucial condition for successful replication at a later date, then, is that *nothing (of significance) would have had to change in the HRM-P system under investigation*. The HRM practices in place in a later period would have to be virtually identical to those in place in a previous period and, furthermore, would have to operate in a virtually identical manner. This is, of course, a most unlikely state of affairs. In another seminal paper in the paradigm, Ichniowski, Shaw and Prennushi[38] consider production lines that 'switch' practices during the period of their analysis. It would be remarkable indeed

---

Replication involves subjecting a theory to the same empirical tests, using the same data source and the same model to see if the original prediction: (a) can be generated again, and (b) remains valid when additional data is added. Reproduction involves subjecting a theory to the same empirical tests, but with a data source and model that can be quite different from the original. Fleetwood investigated the economics of trade unions and concluded that in this paradigm, replication is not undertaken. P. Mirowski and S. Sklivas, 'Why econometricians don't replicate (although they do reproduce)', *Review of Political Economy*, vol. 3, no. 2, 1991; S. Fleetwood, 'The inadequacy of mainstream theories of trade unions', *Labour*, vol. 13, no. 2, 1999, pp. 445-80.

[38] Ichniowski, K. Shaw and Prennushi, 'The effects', p. 302.

equinoxonline

if these lines ceased switching after the period of investigation. Whilst Ichniowski, Shaw and Prennushi are able to identify discrete changes in HRM practices, their meta-theoretical approach leaves them unable to identify subtle, qualitative changes in these practices that occur over time, even perhaps over the time of their study. The way a team operates when it is first set up, is different from the way it operates when it is mature, and even then, it does not continue to operate in this way for ever, it undergoes continual evolution. To suggest, as Godard[39] does, that there is 'evidence that even the most successful programmes may have a limited half-life, either fading over time or failing altogether' is merely to recognise that HRM practices change and evolve: something no-one would deny.

In sum, then, we find no cases of the more discursively powerful notion of prediction_f in the HRM-P research, only cases of prediction_p used to test hypotheses. The workplace appears to be an open system, and in such systems prediction_p cannot be used to test a hypothesis. The scientistic perspective's claim to be superior to other perspectives on the grounds that *it alone can formulate empirically testable predictions* is unsustainable. *The scientistic perspective fails, therefore, to meet its own criteria for scientificity.*

4. *A Concluding Note on Explanation*

Whilst the focus of this paper is on prediction, a few words on explanation are necessary for two reasons. First, because, from the scientistic perspective, prediction is often confused with explanation and this needs to be clarified. Second, because explanation is not a 'poor relation' to prediction. Indeed, if we can actually explain a phenomenon, we are in a strong position to state what kinds of thing this phenomenon might do. This is necessary, although insufficient, for practical or policy intervention aiming to change the world (for the better).

Explanation and statistics

In the lexicon of statistics, to 'explain' is to use the independent (often misleadingly referred to as explanatory) variables to account for some proportion of the variance in the dependent variable. Whilst statisticians are

[39] J. Godard, 'Beyond the high performance paradigm? An analysis of variation in Canadian managerial perceptions of reform programme effectiveness', *British Journal of Industrial Relations*, vol. 39, no. 1, 2001, pp. 28-52, p. 28.

equinoxonline

at liberty to use the term 'explanation' in this very specific sense, it will not satisfy most non-statisticians because it does not explain *why* the independent variables account for some proportion of the variance in the dependent variable. However useful it might be to know *that* X_1X_2, and X_3 'explains' 75% of the variance in Y, neither the equation itself, or the empirical data that constitute the variables, give us any idea why this is the case. A meaningful explanation then, is not simply a matter of 'explaining' some proportion of the variance in a dependent variable.

Explanation is not prediction

From the scientistic perspective, and deriving from the deductive method, prediction and explanation are often conflated in the 'symmetry thesis'. Here, the only difference between explanation and prediction relates to the direction of time. Explanation entails the deduction of an event after it has (or is known to have) occurred. Prediction entails the deduction of an event prior to (knowledge of) its occurrence. If, for example, we can successfully predict that the introduction of a bundle of working practices, when appropriately aligned with corporate strategy, will be followed by an increase in profit, then we can allegedly explain the increase in profitability by the introduction of the HPW practices. Prediction does not, however, constitute explanation. Even in those cases where prediction can be made (almost never in the social world), it is often possible to predict *without explaining anything at all*. Whilst doctors can predict the onset of measles following the emergence of Koplic spots, the latter does not explain measles. An adequate explanation of measles would involve an account of underlying causal mechanisms such as the virus that causes both spots and the illness. Similarly even if we could predict that organisational performance would increase following the introduction of some bundle of HPW practices, the regression equation used to make the prediction would not contain the explanation and we would simply be left asking: Why?

Explanation and tendential prediction

Because of the openness of social systems, events cannot be *inductively* predicted, or predicted as *deductions* from axioms, assumptions and laws—as in the deductive method sketched above. But the social structures, institutions, mechanisms, rules, resources etc. that human agents draw upon in order to initiate action, *can* be *retroduced* and their operation uncovered and *explained*. Retroduction consists

equinoxonline

> in the movement, on the basis of analogy and metaphor amongst other things, from a conception of some phenomenon of interest to a conception of some totally different type of thing, mechanism, structure or condition that, at least in part, is responsible for the given phenomenon. If deduction is illustrated by the move from the general claim that 'all ravens are black' to the particular inference that the next one seen will be black, and induction by the move from the particular observation of numerous black ravens to the general claim that 'all ravens are black', retroductive or abductive reasoning is indicated by a move from the observation of numerous black ravens to a theory of a mechanism intrinsic . . . to ravens which disposes them to be black.[40]

For critical realists, then, retroduction replaces induction and deduction as modes of inference, and explanation replaces prediction as the key objective of science.

To the extent that we can successfully retroduce to the causal structures (etc.) that govern some observation (and there is no gain-saying the difficulty of this), we have a theory that explains this observation.[41] To the extent we have a theory and an explanation, we have an understanding of the tendencies[42] generated by these structures (etc.). To the extent that we understand these tendencies we can make claims about how the structures (etc.) are likely to govern the actions of the human agents that draw upon them. We hesitate to call this a prediction because the term is now so entwined in scientistic discourse that it is almost impossible to untangle it and give it another meaning. Nonetheless, it is a prediction of some kind, albeit heavily qualified, and we call it *tendential prediction*.[43]

How might this work in the case of HRM?[44] If we can successfully retroduce to the social structures (etc.) that, when drawn upon by workers and managers, cause bundles of HR practices to (say) increase organisational performance, then we have a theory with an explanation of organisational

---

[40] Lawson, *Economics and Reality*, p. 24.

[41] On the relation between theory and explanation, see A. Hesketh and S. Fleetwood, 'HRM-performance research: under-theorised and lacking explanatory power', under review, 2006; and D. Whetten, 'What constitutes a theoretical contribution?', *Academy of Management Review*, vol. 14, no. 4, 1989, pp. 490-95.

[42] For an elaboration of tendencies, see S. Fleetwood, 'Causal laws, functional relations and tendencies', *Review of Political Economy*, vol. 13, no. 2, 2001, pp. 201-20.

[43] We wish to thank Anne Fleetwood for suggesting this (critical realist) conception of prediction.

[44] This sketch is, of course, at a very high level of abstraction. The point is to illustrate how we might go about using the meta-theoretical concepts known to critical realism to generate a tendential prediction.

equinoxonline

performance. Such a theory with an explanation would allow us to understand the tendencies generated when workers and managers engage with HR practices and social structures (etc.). If we understand these tendencies we can make *tendential predictions*. We might, for example, be able to understand the tendencies generated by the exercise of human labour power to activate workers' powers for creative, imaginative, ingenious, self motivated and self directed action, as well as the countertendencies generated by the alienation, exploitation and commodification of human labour power. We might, therefore, be able to assess the efficacy of tendencies and countertendencies, and make a *tendential prediction* about the likelihood of specific bundles of HR practices increasing organisational performance.

At present we cannot do any of these things, because we lack plausible theories and, therefore, plausible explanations of organisational performance.[45] And one of the reasons we lack these things is because research is almost totally preoccupied with testing predictions and hypotheses about empirical associations between HRM practices and organisational performance. We feel it is time to consider alternatives.

[45] Fleetwood and Hesketh suggest that current theorising in the HRM-P paradigm is exceptionally weak, a point recognised by some empirical researchers. S. Fleetwood and A. Hesketh, 'Theorising under-theorisation: research on the human resources-performance link', under review, 2006.

equinoxonline