Proposed Book:

**Contributions to
Psychological and Psychometric Theory
arising from studies with
Raven's Progressive Matrices and Vocabulary Scales**

Edited by

John Raven

30 Great King St.,
Edinburgh EH3 6QH
Scotland

Phone: ++ 44 131 556 2912
e-mail (not for publication) jraven@ednet.co.uk

Version Date: 10 June 2006

Overview

Most of the chapters in this book summarise recent research on, or with, Raven's *Progressive Matrices* and *Vocabulary* tests. All deal with significant theoretical or methodological issues which do not, however, relate clearly to one or another of the main preoccupations of modern psychology. Some of the chapters have already been published in journals that are hard to locate; others are only available from Universities as dissertations, but most have not previously been published at all. It was therefore felt that it would be convenient to bring them together in one place.

The book is best understood as a series of reports on journeys of exploration into uncharted territories rather than as a series of classical "scientific" studies with hypotheses derived from previous literature. In the course of each of these journeys "chance" observations often led to the realisation that there were yet other avenues to be explored and hence to the initiation of attempts to find ways of doing so.

Rather than attempt to offer a logical justification for the way in which the chapters have been grouped, the best approach would seem to be for the reader to review the chapter titles and summaries.

General Introduction and Overview: The RPM, its theoretical basis and Measurement Model. John Raven. Filename: RPMBOOK: GI&O.doc.

Linking Psychometric and Cognitive-Developmental Frameworks for Thinking About Intellectual Functioning. Irene Styles and David Andrich. Combined and shortened version available filename: MATRICES: STYLES SUBDIRECTORY: styjr.doc.

By showing that the Item Characteristic Curves for a series of Piagetian tasks map directly onto the ICCs for the SPM items, this paper offers striking evidence for the "existence" of "general cognitive ability" and its measurability via a set of items of increasing complexity and conforming to the requirements of Item Response Theory.

Performance on Raven's Progressive Matrices: What makes a difference: Pictorial Presentation and the effects of timing. Linda Vodegel-Matzen. Status: Chapter 6 edited down awaiting Linda's approval; Chapter 3 scanned in and copy edited, but not edited down. Filenames: RPMBOOK: Linda6 and Linda3.

One of these studies shows that, when the items are re-constructed using concrete objects (hats, bananas etc) instead of abstract symbols (but in such a way as to preserve their logic), all the items become easier for everyone. The items retain their order of difficulty and, most importantly, the change does not differentially benefit the "disadvantaged". In this (and other) ways the dissertation strongly reinforces the scientific status of "general cognitive ability" – it is not "merely a convenient statistical construction" -- and its measurement using the RPM. The other shows that more able respondents work more *slowly*.

PART II

PRACTICAL MEASUREMENT ISSUES:
LESSONS FROM 75 YEARS' WORK WITH *ITEM RESPONSE THEORY*:
BENFITS, PROBLEMS, AND POTENTIAL SOLUTIONS

Section Introduction: John Raven. Filename: IP2.doc. Status: available.

In developing his *Progressive Matrices* tests, J.C.Raven anticipated the development of Item Response Theory in that he plotted what have since become known as "Item Characteristic Curves", plotting the curves for all items on the same Figure. These he examined for abnormal trace lines (which suggested that there was something wrong with the items), cross-overs, and equality of spacing.

Although the measurement model he employed was briefly explained in the "Guides to the use of" the various *Progressive Matrices* tests that were published at the time and in a number of journal articles, the ways in which it differed from Classical Test theory were not made explicit. It was therefore difficult for

3

researchers, without the benefit of more recent formulations of IRT, to appreciate the implications and the tests continued to be "misused" and the results misinterpreted.

Only recently, by, with considerable difficulty, replicating Raven's methods using modern IRT programmes has it become possible to appreciate how close Raven had come to placing the scientific status of "general cognitive ability" – or, more accurately, "eductive" ability - beyond dispute, and the RPM as a measure of it.

The chapters in this section belatedly rectify these oversights.

The Need for, and Development of, the SPM+. John Raven. Status: Available. Filename: RPMBOOK: SPM+.doc

So far .doc versions of the necessary Figures have not been located, but I think they came back from Andy Corsham to Jean. The Oxford Illustrators versions are in the CRC file behind my desk.

The Romanian Standardisation of the SPM+: Sample and General Results. Anca Dobrean. Filename: RPMARTS:ROMSTMJS.doc or ROMSTMJRWPE.pdf. Available on WPE website.

In the winter of 2002-03 a nationally representative sample of 2755 Romanians aged 6 to 80 were tested with the *Standard Progressive Matrices Plus* in their own homes. This chapter describes the sample and summarises the available information on the ways in which the scores varied with socioeconomic and demographic variables. The other chapters in this Section summarise major statistical studies carried out using these data.

Using the Romanian data to Replicate the IRT-based Item Analysis of the SPM+: Striking Achievements, Pitfalls, and Lessons. John Raven, Joerg Prieler, and Michael Benesch. Filename: RPMARTS: CV&IRTDIR: CV&IRTPS.doc + CV&IFIG.doc or CV&IRTWPE.pdf. Electronic version available on WPE website.

Only recently has it become possible to plot, and then only via labour-intensive customised adaptations of the standard programs, sets of ICCs derived from "sophisticated" Item Response Theory (IRT) programs in a form analogous to the multi-item plots Raven published in the mid 1930s. In order to compare these with each other and with the results of applying Raven's original techniques, Gerhard Fischer developed a computer program which replicated Raven's graphical methodology.

When the Raven-type "empirical" graphs derived from the Romanian data were compared with computer-drawn curves generated by fitting 1 and 3-parameter IRT models to the data, it turned out that adoption of the most frequently used version of IRT – the 1 parameter model – can induce test developers to make serious errors. On the other hand, those produced using the 3-parameter model give a reasonable approximation to the "empirical" graphs.

4

But the most striking result of the study is that the plot of the ICCs for all 60 items strongly confirms the scientific existence of "eductive ability" – it is not merely a statistically based device for organising data - and its measurement using the RPM.

Lessons Learned while Developing a Romanian Version of the Mill Hill Vocabulary Test. John Raven, Camelia Rusu and Michael Benesch. Current summary version, based on a small preliminary data set by John and Jean Raven only. Filenames: MATRICES:ITC06:LFROMHV.doc; Tables in ITC06:RUMTAB.xls Figures in ITC06:RIDCC.doc. Requires final versions of TCCs and TIFs.

This chapter discusses problems arising while developing a Romanian version of the *Mill Hill Vocabulary* scale, together with the steps taken to overcome them. Most versions of the *MHV* are made up of two sets of 34 words with one Set in Open Ended format and the other in Multiple Choice format. The question then arises of whether the words are easier when presented in multiple- choice format and whether they scale in the same way.

It is not so easy to check on this as might be assumed because a 2-parameter model is appropriate when the test is administered in open-ended format. On the other hand, when they are administered in multiple-choice format allowance has to be made for "guessing". It is therefore necessary to employ a 3-parameter model. Unfortunately, one is then left with the question of how comparable are the indices derived from the two models.

It is evident from the Distractor Characteristic Curves that were prepared that some distractors seriously confuse people who, in reality, know the correct answer.

But this is a high level problem compared with the technical difficulties actually encountered when it comes to trying to generate sets of ICCs derived from the 3-parameter analyses in the first place. IRT indices are very unstable unless they are derived from very large samples – and this problem becomes even more serious when 3-parameter models are employed. Yet, one cannot use large, nationally representative, samples when *developing* a test. Such data only become available in the standardisation phase. Consequently, one can expect to get different results when the data obtained in these two phases are compared.

Problems in the Measurement of Change (with particular reference to Individual change [gain]Scores) and their potential solution using IRT. Joerg Prieler and John Raven. Filename: MATRICES: RPMARTS: MCGIDIR: MCG31APA.doc and MCGWPE.pdf. Electronic version available at WPE website.

Part I of this chapter reviews some major, but commonly overlooked, problems in the measurement of change and, in particular, in the calculation of change scores (such as before vs. after difference scores purporting to index variance in people's responsiveness to such things as instruction, stress, therapy, or drugs). One of the most important is that the raw score differences that correspond to equal differences in latent ability vary markedly with the absolute difficulty of the test employed, the shape of its test characteristic curve, and the sector of that curve on which the change is measured. This is true even on tests which satisfy the requirements of the most popular versions of Item Response Theory.

In Part II, an IRT-based methodology which has been developed by Gerhard Fischer to overcome this problem is described and its use illustrated.

Establishing uni-dimensionality in IRT-based Tests: David Andrich. Status: TBW

Although J.C.Raven and Rasch (and perhaps Guttman) both illustrated the way in which sub-sets of ICCs sometimes behaved in ways which were similar to each other but different from those of other items in the test and how this fact could be used to demonstrate that they belonged to a different dimension, most IRT theorists have continued to discuss the problem of establishing unidimensionality from the viewpoint of classical test theory. This chapter presents a new, theoretically-based, perspective on how the problem is to be addressed and illustrates its application through yet another re-analysis of the Romanian data.

PART III

STABILITY AND CHANGE IN RPM NORMS ACROSS TIME AND CULTURES
... and their implications.

Stability and change in norms over time and Culture: The Story at the Turn of the Century. John Raven. Filename: SAC. doc Requires checking of fig nos etc. and references. Figures not available in electronic form. Earlier, Cognitive Psychology, version available in hard copy only.

Data relating to the stability and variation in the norms for the *Raven's Progressive Matrices* Test for different cultural, ethnic, and socioeconomic groups on a worldwide and within-country basis are first summarized. Subsequent sections deal with variation over time. The inter-generational increase in levels of eductive ability over time created a need to restore the test's ability to discriminate among more able adolescents and young adults. A version of the test – the SPM *Plus* - which overcomes this problem without reducing its discriminative power among the less able is discussed. In addition to offering a possible explanation for the variation in norms over time and between ethnic groups within countries is offered, the chapter reviews research into the educational and cultural variables known to affect RPM scores.

The SPM+ in Hungary. Balazs Klein, Sandor Klein, Kalman Joubert, and Gyula Gyensis. Status: Paper available in Hungarian. Needs translation into English.

In 1998 more than eight thousand 18 year old Hungarian conscripts filled out a questionnaire developed by the Central Institute of Statistics. Of these, almost 7 thousand also completed the *Standard Progressive Matrices Plus.*

This chapter discusses the way in which SPM+ scores varied with background variables. These included information on parents' level of education on the one hand and conscript's level of education, school grades, area of residence, height (at birth and at conscription), hygiene and nutritional habits, habits such as smoking and drug taking, and the number of books and computers in the home.

The chapter also discusses the interactions between the background variables and the way in which they have been composited to yield a measure of socio-economic status.

More recent comparative data for the SPM across cultures: Kuwaiti norms for the Classic SPM in an International Context. Ahmed Abdel-Khalek and John Raven. Status: hard copy published version available. Filename: MATRICES: RPMARTS: KWAAKJR.doc

A probability sample (N=6,594) of Kuwaiti school students aged 8-15 responded to the *Standard Progressive Matrices*. The test was administered, untimed, in group sessions. In this chapter, the smoothed summary age norms for Kuwait (which will themselves be of interest to many psychologists and others working in Kuwait and neighbouring countries) are first compared with what has become the standard international reference data for such work, namely the 1979 British norms, and thereafter with data collected in a wide range of cultures. It emerges that, at any point in time, the norms are remarkably stable across cultures, but have changed dramatically over time. These findings show that, while as yet unidentified features of the environment have a dramatic effect on scores, aspects of the environment that many people would have expected to have a significant effect (such as differences in calligraphy) are much less important than might have been expected.

Overview of Cross-cultural Norms for the SPM+. Anca Dobrean. Status: to be written.

In this chapter, the Romanian data reported earlier will be compared with similar data from Hungary, Germany, Poland and the USA. Particular attention will focus on the problems posed by the varying sampling procedures used in the different studies.

The chapter will include a discussion of the way in which the term "sample" is widely mis-used by psychologists, who frequently do such things as apply significance testing in situations where it is entirely inappropriate. For example, many psychologists apply significance testing to sex differences within groups which can by no stretch of the imagination be regarded as samples of men and women and then draw apparently generalized conclusions.

Lithuanian Norms for the Coloured Progressive Matrices in a Cross-Cultural context. Grazina Gintilene and Dovile Butkiene. Lithuanian version available in hard copy: requires translation into English.

This chapter reports a standardisation of the *Coloured Progressive Matrices* (CPM) among Lithuanian children. A representative sample of 1067 children aged six to eleven was tested after first selecting a stratified sample of schools/kindergartens by place of residence and language of instruction. Within schools, sampling was carried out randomly within age and gender groups. Children receiving special education services within mainstream schools were included. The sample for age 6 children included those not attending school as well as those at school or kindergarten. Social demographic and academic information

on the children was gathered using a parental questionnaire. Data analysis based on methods of internal consistency, item analysis, and comparison of data between different groups of children showed adequate psychometric properties of the CPM. The data also indicate that it is important to consider variables in the home and school environments when interpreting the results of individual testing.

A pilot study to establish xhosa-speaking norms for the Coloured Progressive Matrices. Natalie Bass and Jan Knoetze, Dept. Psychology, Rhodes Univ., Grahamstown, South Africa 6140. Shortened version, not yet approved by Natalie, RPMBOOK: BASSJR.doc. Electronic copy of published paper: MISPMST DIR: BASSSACPM.doc.

The Coloured Progressive Matrices is used extensively across a wide variety of settings in South Africa. However appropriate local normative data has yet to be collected. This pilot study thus sought to establish normative data for a population of Xhosa-speaking Primary School children in the peri-urban township area in the Grahamstown region. The booklet version of the test was administered in group format using an alternate method of test administration (involving Xhosa instructions) developed by Vass in 1992. The sample consisted of 197 male and 182 female Xhosa-speaking children in Grades Two to Seven (N=379).

The results showed: (1) a significant effect of age on test scores, where scores increased with age as expected; (2) no significant effect of education on test scores; and (3) that the norms were in general lower than those obtained by Raven, Court and Raven (1990) during the standardisation of this instrument in the United Kingdom and America.

It was concluded that (1) there is an urgent need for more appropriate South African normative data for this test; and (2) that when assessing African children from disadvantaged backgrounds, further research into the effects of cultural and socio-economic factors and gender on non-verbal intelligence (and on performance on this test in particular) is required.

The Coloured Progressive Matrices in South Africa. Adien Linstrom, John Raven and Jean Raven in collaboration with Jopie van Rooyen and Ptnrs. RPMBK: SAMCK.doc. WPE Version: RPMARTS: SAWPE.pdf. Previous version with suggestions for further developments:RPMBK: SA.doc + separate tables

The nature of the challenges faced by this country is reflected in the fact that it has eleven official languages and a population of 44,700,000 divided into four main groups. Black people represent a diversity of indigenous groups; white people are mainly descendents of European immigrants; Coloured people are descendents of cross-cultural relationships, and Asian people are mainly Indian. Included are also minority groups of Chinese, Taiwanese and Japanese. The population mix is unique in that Whites and Blacks have their roots in two totally different worlds. The first is primarily a European capitalistic industrialized society and the second mainly a pre-industrialized way of life.

Of the total population only an estimated 6% of adults have any post-school qualifications, 16% have school-leaving certificates and only 25% have completed primary school. There is an urgent need to focus on strategies to accelerate the education and development of the youth of South Africa. With this in mind, the Education Department of the Free State requested assistance in establishing norms for the Coloured Progressive Matrices.

Completed CPM profiles were obtained for 2,460 South African children between the ages of 5 and 12 years. The results reflect responses from different language groups and economic environments, private, government and farm schools, and different home environments.

The SPM among Slovakian Gypsies. Jan Ferjencik. Staus: 2 hard copy unpublished papers in RPM box. Need to be scanned in and edited. Jan has several other papers and analyses in incomplete form.

Peruvian Mountain Norms for the Classic SPM. John Raven. Status: Only norms available. Sample and Discussion of eg effects of altitude needs to be written.

Tribal vs Urban norms for the Classic SPM in India. C.J.Deshpande. Status: Tribal data to be collected. Promised by June 2006.

Urban norms were collected from representative samples of more than 5,000 young people in both Pune and Mumbai (Bombay). Contrary from what might have been inferred from earlier studies based on less representative samples, these Indian norms proved to be remarkably similar to may now be appropriately termed the international norms for cultures with a tradition of literacy.

PART IV:

OUTSTANDING ETHICAL, CONCEPTUAL, AND MEASUREMENT ISSUES

Searching for Justice: The Discovery of IQ Gains over Time (and the Pursuit of Humane Ideals.) Jim Flynn. Status: Flynn, J. R. "Searching for justice: The discovery of IQ gains over time" was published in the *American Psychologist*, 1999, *54(1)*, 5-20. *How to Defend Humane Ideals* (2000) is a huge book. Extracts from the latter are being consolidated in   . A summary of *Asians Americans: Achievement beyond IQ*, which dramatically demonstrates the effects of many variables neglected by psychologists is available in RPMBOOKDIR: FLABIQ.doc.

A Response to Flynn John Raven. Status: available: filename RPMARTS: RTF2.doc and RTFWPE.doc. Available at WPE website.

Jim's article opens with an attack on the validity of the RPM and the way in which scores have been used to legitimize the assignment of second-rate status to certain ethnic groups. It continues with a demonstration that other important variables – such as creativity, motivation, and social support – have been neglected. But it then launches into a discussion of meritocracy. And here is the real problem with

the article - because Jim fails to discuss either the legitimacy of the concept of hierarchy into which notions of meritocracy fit or the forces which support hierarchy and require a single-factor concept of ability to legitimize the unjust discriminations Jim wishes to ameliorate. Those discriminations compel participation in a system which is not merely inhumane but is also relentlessly driving *homo sapiens* toward both its own extermination as a species carrying the planet as we know it with it.

My response is divided into two main parts. The first deals with Jim's attempt to discredit the meaningfulness of the *Raven's Progressive Matrices.* The second discusses five other issues pertinent to the advancement of humane ideals through psychology.

Intelligence, Engineered Invisibility, and the Destruction of Life on Earth. John Raven. Status: available. Filename: Doctemp: EI12APA.doc. Also several BJP versions. WPE version: EI12WPE.pdf

Preparation of this paper was precipitated by the contrast between the conclusions drawn by Gottfredson (1997) and those emerging from papers brought together by Raven and Stephenson (2001). Gottfredson assembled data supporting three main claims: (i) g is the main variable responsible for most of the variance in occupational performance; (ii) g is the most important variable determining the effectiveness of behaviour outside work; and (iii) occupational status depends mainly on g. In this paper it is shown that, in both the workplace and the educational system, other qualities *are* important but remain invisible. This invisibility stems in part from an inappropriate psychometric model and limited criteria of performance but mainly from a societal need for a single and unarguable criterion of merit to legitimise a social hierarchy which compels most people to contribute to the unethical activities of which modern society is so largely composed. Embracing the task of mapping these socio-cybernetic forces results in focussing on the external rather than the internal determinants of behaviour – in effect, turning psychology inside out. More specifically, understanding them enables us to specify the arrangements required to run the educational system – and many other domains of human endeavor – effectively. These developments depend quintessentially on organisational arrangements, job descriptions, and appraisal systems to be introduced by psychologists.

Psychometrics, cognitive ability, and occupational performance. John Raven. Status: available. Filenames: RPMARTS: PCA&OPI.doc and PCA&OPIR.doc. Requires both files; the latter contains the references for the first.

In two earlier articles (Raven, 1989b, 2000), I reviewed studies suggesting that, contrary to what Flynn (1987) would have us believe, the *Raven Progressive Matrices* measures psychological abilities of fundamental importance and that the steadiness in the improvement in these abilities over time and the similarity in the norms obtained in many - but not all - cultures at any point in time reinforce this conclusion.

In this article I summarise remarkable new evidence that the *Raven Progressive Matrices* is measuring an important aspect of cognitive functioning. Thereafter, I return to the question of the extent to which it measures "intelligence" (and

10

competence more generally). This leads to a re-examination of the test's construct validity. This discussion has important practical implications because it underlines the need to situate eductive ability scores in the context of a yet-to-be-developed framework for thinking about the wider aspects of intelligence and competence. At the same time, it raises serious questions about the way we think about the procedures to be used to establish the validity of a test and the ethics of insufficiently *comprehensive* assessment - however *in*valid some of the necessary assessments may be. The article concludes by outlining some of the parameters which must be satisfied in seeking to develop a better framework for thinking about competence and its assessment.

PART V

EMERGING APPLICATIONS

Predicting Driver Disqualification. Thomas Karner, Dept. Psychology, University of Vienna. Status: Abstract is in RPMARTS: VTSDRIVACCDNS.doc. Karner left Schuhfried under a cloud and it will be necessary to get the full info from Joerg. Somewhere we already have a table of predictive validities.

The validity of the Traffic Psychology Test Battery of the "Vienna Test System" in predicting driver history (n=1095) – and especially loss of license to drive – and the correlation to real driving behaviour (n=80) were assessed. The Traffic Psychology Test Battery as part of the VTS includes measures of reactive capacity under conditions of stress, visual structuring ability, concentration, attention, reaction time and many aspects of personality as well as the *Standard Progressive Matrices*. Given the range of tests included, the SPM had surprisingly high predictive validity.

The Detection of Malingering: McKinzey, Raven & Prieler (2 short papers). Status: Available: Filenames: RPMARTS: MCKRAV.pdf and MCKIDSO3.pdf.

McKinzey, R. K., Podd, M. H., Krehbiel, M. A., & Raven, J. (1999). Detection of Malingering on the Raven Progressive Matrices: A Cross-validation. British Journal of Clinical Psychology, 38(3), 435-439.

A formula for detecting faked Raven profiles was cross-validated on 48 experimental malingerers and 381 people from the standardization sample. The formula yielded a cross-validated 29% false negative rate and a 9% false positive rate.

McKinzey, R. K., Prieler, J., & Raven, J. (2003). Detection of children's malingering on Raven's Standard Progressive Matrices. *British Journal of Clinical Psychology*, 42(1), 95-99.

A formula for detecting faked Raven's SPM profiles was cross-validated on 44 children and adolescents (ages 7-17). It yielded a false negative rate of 64%. However, a formula involving three very easy items (i.e., any of A3, A4, & B1 missed) yielded a hit rate of 95%, with 5% false positive and negative rates. All but

two of the participants were able to produce lower scores when asked to fake the test.

Too Dumb to Die Kim McKinzey. (This whole chapter is now in RPMBOOK Dir. Filename: TDTD.pdf [because it's a .pdf file it doesn't open in Word, but has to be opened from Windows explorer]. However it has also been exported as a .txt file from which this abstract has been extracted.)

In Atkins v. Virginia (2002), the US Supreme Court held that executing the mentally retarded is unconstitutional. In a capital, death penalty case, a hearing must therefore be held sometime before sentencing or trial to determine whether or not the defendant is mentally retarded. An Atkins case study is presented, wherein the issues involved are discussed. These issues include: timing of the hearing, burden of proof, malingering, data gathering, and measurements of intelligence and adaptivity. [One key question is whether or not the defendant's IQ is above or below 80(?) which raises the whole question of dates of norms, statistical procedures used to generate norms in the tails, discriminative power of tests, equivalence of cut off scores across tests discriminative power of tests, and role of psychologists. JR]

Lethal Intelligence. Raven. Status: Available: filename: Doctemp: TDON.doc

This paper is based on two extraordinary documents circulated by Kim McKinzey (see tinyurl.com/4mp3r and tinyurl.com/3opys). The issue turns on the fact that, crudely over-stated, the laws relating to the death penalty for murder in some States allow that the actions of mentally retarded murderers can be excused because they are likely to have failed to understand the implications of their actions.

So how to determine whether someone is mentally retarded or not? What we see in the second of the documents McKinzey circulated is that a host of tests, ranging from the Wechsler Intelligence Tests to the Vineland Social Maturity scales, have been deployed by forensic Psychologists, their relevance disputed, and the results compared with "common sense" assessments of "real life" behaviour.

But behind such antics lies another set of disputes: How good are the samples on which the norms are based? If prospective participants in such studies have exercised their "informed consent" based rights not to participate, what effect has that had on the norms? What is the effect of the *date* on which the norms were collected? (Judged against yesterday's norms one should die; yet, given today's norms, one may live). And what statistical procedures have been deployed to compile the norms? (As Dockrell[1] has shown, the IQ of the same person on the same test judged against the same norming sample can vary dramatically depending purely on the assumptions made by the statistician who processed the data.)

Yet, to me, even such questions miss the point. That point has two facets – one to do with ethics; the other with competence.

One of the most surprising conclusions to emerge from both our own work and that of others (such as Donald Schon), is that incompetence in modern society stems above all from an inability and unwillingness to engage with the wider social forces which *primarily* determine behaviour and thus what people *can* do in their jobs. In this case, this implies that we, as psychologists, need to get together with others (perhaps through our professional organisations) to influence the social and legal contexts in which we work instead of accepting those contexts as givens. Furthermore, unless we do this we *cannot* behave ethically.

Legitimising Payments for Remedial Education and Health Care: The requirement to demonstrate individual and group benefit: German health care; US insurance. Ralf Horn has agreed to contribute. Tommy McKay's later work outside this area so dropped as a potential contributor.