

Editorial

Welcome to the fourth issue of *Building Research Capacity*. Since the last issue the RCBN have put together an outline of activities we are proposing to organise over the next two years. All these activities, ranging from one-off events to on-going discussions, have been organised around particular methodological and career development themes (see the back page for more information). The four main articles in this issue each relate to a different one of these RCBN themes. One such theme is the nature and use of *combining methods* in teaching and learning research. The article by Nash (p.1) presents a realist approach to social explanation and argues that a 'numbers and narratives' methodology offers a way for combining data. The second article by Bradshaw (p.4) discusses the use of contracts in research and argues that their use can inform and empower the researcher and the participants. This, he argues, would add value to the use of 'qualitative' *in-depth data-sets*, another of the RCBN themes. The third article by Coe (p.6) briefly outlines what an effect size is, how it is calculated and how to interpret it. Effect sizes are further explored and considered in the RCBN theme *the role of numbers* in teaching and learning research. The final article by Lambert (p.9) is closely related to a further theme of the RCBN's activities, *the use of secondary data* in teaching and learning research. In this article Lambert discusses practical issues in handling occupational information.

A realist scheme for social explanation: on 'numbers and narratives'

Roy Nash

Massey University College of Education
New Zealand

Sociological explanation and some difficulties with the standard model

It must go without saying that the object of social research is to generate explanations of social events and processes considered in some way problematic by researchers and those who fund their activities. As this is almost self-evident one might think that arguments about the nature of successful explanation in social science would be long settled. Such, however, is not the case. The outline presented here of a realist approach to social explanation may, in this context, thus serve as an introduction to an unduly neglected aspect of social research (Nash, 2002a).

One form of explanation held to be definitive of sociology attempts to account for social facts in terms of other social facts. This might be called the level of system-to-system explanation. Most sociological accounts of this kind attempt to link two states of affairs, Bunge (1998) calls them "function states", which are represented by indicator variables. There is no shortage of examples of such research, but a US study showing a negative relationship between the rate of tax levied on a six-pack of beer and the rate of sexually transmitted diseases among young people, has the advantage of being particularly memorable (Dominion, 2000). In this in-

stance, the price of beer (function state one), relaxation of inhibitions (individual dispositions), sexual activity (social practice), and rate of sexually transmitted diseases (function state two) are evidently related. In Bunge's view, such states of affairs exist (when they do) and, because they affect what people do and can do, they may be studied by realist social science and included in causal accounts. It is often possible to construct equations where the relationship is modelled by the form, "for each unit of increase in variable *a* there is observed a decrease of *y* units in variable *b*", which sometimes have a high degree of predictive efficiency (apparently, a 20% increase in the tax on beer is associated with a 9% reduction in STDs). Models of this kind give explanations acceptable for certain purposes, but their limitations should be apparent. The most critical is the omission of an effective causal mechanism which must include reference to the dispositions of agents and the their concrete practices which generate the observed effect for which an explanation is sought. We only understand the relationship observed because of the taken-for-granted knowledge we have about adolescent sexual practices.

A related difficulty with such accounts lies in the interpretation of the "at risk" models that now provide

(Continued on page 2)

(Continued from page 1)

a standard form of analysis and explanation in policy-making circles. A kind of shorthand has emerged in which behaviour is increasingly explained by "risk factors", and while there is a practical utility to this, there are definite limits to the form of explanation offered. To argue, for example, that young women from specified social class and ethnic groups are over-represented among those who become pregnant as teenagers because the empirically proven risk factors are more common among their families, is little more than vacuous (Nash, 2000). As an explanation of a social practice, to say that those who adopt it do so because they are the kind of people who probably will do so, certainly takes one no great distance towards understanding *why* recognisable forms of practice exist, *why these* individuals (rather than others with similar "risk" characteristics) are affected, or *why* they number 10% or 20%, and not some other figure. The related notion that all students from a particular group, identified, for example, by its social or ethnic origin, are each equally "carriers" of a specified weight of disadvantage, a virtual handicap, is a particularly common error. As Humphreys (1984: 2) points out "[c]hance by itself has no causal properties or powers". These persistent errors, however, have a long history and are difficult to resolve within the nominalist epistemology that dominates so much quantitative research. They do not, of course, negate the value of quantitative investigations, but they would be better corrected. The way forward, in fact, is not to abandon quantitative analysis for qualitative methods but to seek to integrate both in a realist scheme.

A realist concept of the social

In a realist philosophy the task of physical science is to discover the essential structures of natural things and specify the mechanisms by which effects observable at the level of appearances are generated. The same principle can be applied to the study of human societies. Bhaskar's (1993) naturalism thus constructs

the tasks of physical and natural science as being essentially parallel: both are concerned to investigate the nature of their object and to describe the mechanisms by which their particular properties and effects are produced. Of course, the procedures for demonstrating the existence of social forms, and demonstrating that certain observed phenomena are effects of those forms, can be peculiarly difficult. To say that social forms instantiated in practice are real does not make their nature and properties any easier to demonstrate, nevertheless, although the different ontology of social and physical entities set limits to the possibilities of naturalism, those possibilities do exist.

In the first place, we should regard social organisations as real entities with emergent properties. In this way, families, firms, football teams, and so on, are held to be entities with emergent properties appropriate to the form of their organisation. A family is thus constituted by a specific set of social relations which, being organised to form a social system, constitute its *structure in that respect*. A family then possesses resources (including the embodied *dispositions* of its members) necessary to the accomplishment of its social functions and, in the performance of those tasks, adopts distinctive social *practices integral to those functions*. In other words, *families are* social organisations with certain definitive social functions (particularly those related to rearing the young), and can be recognised as groups of people united by socially recognised bonds, equipped with the material and symbolic resources essential to their collective purposes, and necessarily engaged in the social activities through which those are achieved. The analysis of social entities and their distinctive properties requires, if this analysis is correct, recognition of the structures, dispositions, and practices that give social entities their specific character. A realist explanation of social events and processes will also have exactly this structure-disposition-practice (SDP) scheme. In the sociology of education, where this formulation was de-

veloped, social differences in access to education, for example, can be explained in terms of the social structures, socialised dispositions, and social practices of – at least – families, schools, and firms (Nash, 2002b). With respect to the family, for example, an integrated explanatory narrative might include reference to material resources (particularly those largely given by *structural* class-position such as wealth, education, and social connections), psychological *dispositions (such as parental aspirations)*, and *recognised practices* (everything from bedtime stories to advice on university application), which are all at their different levels effective in "making the difference".

These observations will sound commonplace, and they are certainly not intended to offend common-sense, for although it may be an error to celebrate common-sense it is perhaps an even greater error to dismiss its insights. As for social research, it will be apparent that any systematic investigation of, for example, the family should be concerned with (i) its structure (the character of the relations that constitute it together with its dedicated resources); (ii) with the dispositions of its members that enable them to act as husbands, wives, parents and children according to the norms of the society that confers on such relationships their nature and powers, and (iii) the established practices within which such acts are performed, for it is through practices that social organisations have their agency. In as much that this explanatory scheme recognises the level of dispositions, as habits of mind and body acquired through socialisation, it calls for a co-operative relationship with psychology. There is no need for psychology and sociology to compete for the proportion of variance each can explain.

The SDP scheme outlined here is not offered as anything new. All attempts to develop an integrated sociology have recognised that the explanation of social events and processes must include reference to

(Continued on page 3)

(Continued from page 2)

social structures, socialised agents, and social action or behaviour. The scheme might, in fact, be seen as a realist and "anglicised" appropriation of Bourdieu's sociology (Bourdieu, 2000). There are, however, some implications for the integration of quantitative and qualitative research that merit comment. A realist approach to the construction of quantitative models requires that the properties of social entities be expressed in an appropriate concept and their quantification effected by indicators adequately supported by theory. In the sociology of education, for example, it is necessary to construct indicators of structural properties (basically material resources), dispositions (basically symbolic properties), and established practices. When variables have been constructed with a clear understanding of the properties they are supposed to indicate it is likely that the capacity of the model to represent the reality of the complex social processes it seeks to explain will be enhanced. In the sociology of education, for example, it must be an error to include indicators with a structural reference (e.g. number of books in the home), a dispositional reference (e.g. parental aspirations), and a practice reference (e.g. bedtime stories) in a composite index of "home quality", when the properties these indicators point to are not *additive* but related at different levels to the family as a social organisation.

The standard quantitative research programme is thus open to a realist critique on several grounds. One of the most difficult concerns the nature of measurement itself in social science. This is a matter of great philosophical complexity, but it is worth mentioning that the quantification of extra-physical properties is generally effected by methods of scaling that cannot approach the level of measurement precision achieved in physical science. What this means for practice is very much a matter to be taken up in the context of specific studies, but there are implications for social science, and particularly for the twin legitimating

concepts of "reliability" and "validity" that advanced researchers might find worth the effort of study (Berka, 1983).

Conclusion

Scientific realism is committed to the belief that social events and processes can be observed and their several causes determined by controlled observations and experiments. The realist concept of causality in this context is worth a brief comment. Realism explicitly rejects the positivist doctrines of (Humean) causality, "predictive" explanation, conceptual nominalism, and measurement operationalism that continue to underpin much, although not all, quantitative research. Realism does not, however, hold any prejudice against quantification, endorses moderate empiricism, accepts in most versions a correspondence theory of truth, and argues for methodological rigour in the search for effective social mechanisms. There can certainly be no support in a realist social science for the attempt to win greater privilege for data obtained by one method of collection rather than another, but in as much as the structure-disposition-practice scheme calls for the use of complementary methods it is not opposed to ethnographic or "qualitative" research. Indeed, it is only through such methods that the study of *structures, dispositions, and practices* can be achieved. The special role of quantification is to enable the relative contribution of indicators, and therefore the properties they represent, to be weighed.

There is no dispute that quantitative methodology can provide important information about sources of variance in the distribution of social goods and, through multivariate analysis, can provide an investigator with a shrewd idea of the various social processes that have generated them. It remains true, however, that the interpretation of statistical patterns without any direct knowledge of the social processes by which they were generated does not provide a deep explanation. Although it is not always easy to identify the social processes responsible for effects detected at the system

level, an integrated realist approach is more likely to disclose what is happening than research supported by a taken-for-granted positivism. Realists are also conscious that the rejection of nominalism means that hermeneutic skills, on which qualitative researchers pride themselves, are actually critical to the interpretation of variables and their pattern of interaction in quantitative models. Statistical models that include the levels of system properties (structure), individual dispositions, and social practice, with indicators appropriate to each level, can be difficult to interpret without a strong theoretical basis. In fact, at the risk of sounding unduly gnomish, one might say that realism requires the hermeneutic enterprise to be resisted at the heart of the qualitative project, where it can interrogate the construction and interpretation of indicator variables (Nash, 2002c). A "numbers and narratives" methodology able to unite qualitative and quantitative techniques may have something real to offer contemporary social researchers.

References

- Berka, K. (1983). *Measurement: Its concepts, theories and problems*. Dordrecht, Holland: Reidel.
- Bhaskar, R. (1993). *Dialectic: The pulse of freedom*. London: Verso.
- Bourdieu, P. (2000). *Pascalian meditations*, trans. by R. Nice. Cambridge: Polity Press.
- Bunge, M. (1998). *Social science under debate: A philosophical perspective*. Toronto: University of Toronto Press.
- Dominion* (2000). The dearer the beer, the safer the sex, study shows, 29 April, 5.
- Humphreys, P. (1984). *The chances of explanation: Causal explanation in the social, medical and physical sciences*. Princeton, NJ: Princeton University Press.
- Nash, R. (2001). Teenage pregnancy: The case for "numbers and narratives" in policy research, *Social Policy Journal of New Zealand*, 17, 200-13.
- Nash, R. (2002a). Numbers and narratives: Further reflections in the sociology of education, *British Journal of Sociology of Education*, 23(3), 397-412.
- Nash, R. (2002b). A realist framework for the sociology of education: Thinking with Bourdieu, *Educational Philosophy and Theory*, 34(3), 273-288.
- Nash, R. (2002c). Peer effects and peer

(Continued on page 4)

group processes: joining the conversation on qualitative and quantitative approaches, *Australian Journal of Education*, 46(1), 6-18.

Contracts with research participants: degrees of appropriateness

Matt Bradshaw

University of Tasmania, Australia

Increasingly, social scientists, especially qualitative researchers, are including, or are being required to include, contracts with participants in the methodologies of their research. It may be of comfort for those new to contracts to know that what is often a recent requirement of various forms of ethics committees has a somewhat longer history of application. Of particular relevance to readers of this journal is that a considerable part of this history is in education research.

In this article I first introduce contracts. Broadly, contracts have been used to cover the gamut of reporting positions in social research, i.e. from projects studying things which a researcher may wish to 'speak against' to projects involving things which a researcher may wish to 'speak for'. There are different implications for contracts depending on which type of project is being undertaken. Second, therefore, I discuss some of the advantages of contracts and some of the circumstances in which they are appropriate. Third, I discuss their disadvantages and possibly less appropriate circumstances, from which I conclude that contracts can be a useful part of a project's methodology but also need to be approached with care regarding their appropriateness to a particular piece of research.

What are contracts?

'It is increasingly the case that "contracts" with groups and communities [whose members] agree to be researched include feedback of some sort ... These "contracts" are now often demanded by ethics com-

mittees that sanction research on human subjects' (Baxter and Eyles 1997: 515). Contracts require that certain criteria, such as securing the informed consent of research participants, be met before research can proceed (e.g. Babbie 1998). However, and discussed further below, this requirement is not without its difficulties.

In sociology, contracts have been developed and applied for some time in ethnography – particularly regarding research on education (e.g. Ellen 1984; Punch 1986; Wittrock 1986). It is now well accepted in sociology that research participants need to be protected, and many standard contract documents are available to present to would-be participants informing them of their rights and requiring their consent before they can be involved in research (e.g. Berg 1998). One right commonly extended as part of a contract is that participants be given the opportunity to check the use of their material in any reporting document. 'Member checks', as they are often termed, have been most influentially recommended by Guba and Lincoln (1981; 1989). Member checks are argued to be a key to establishing the accuracy and credibility of many types of qualitative research.

When are contracts appropriate?

A distinction is made in the literature between researching 'the powerful', or 'up', and researching 'down' (e.g. Cormode and Hughes 1999). Researchers study groups subalternated by dint of sex, race, ethnicity, religion, age and so on. As well as being disadvantaged, such groups

are considered to be deficient in power and/or resources in some way. An advocacy role for the researcher may develop in such instances, and a critical position may be taken against structures operating to disadvantage such groups.

The researcher, too, occupies a different as well as privileged social and speaking position from that of his or her research participants, and many researchers are concerned in some way to redress this imbalance. Negotiating a clear set of rights and undertakings between researcher and participants is seen as one way to empower participants. Moreover, research participants have rights that researchers have a responsibility to respect. Chief among these rights/responsibilities is that research participants do not come to any harm through being involved in a project. Many qualitative researchers go further than this, especially when studying 'down', considering that the relationship between researcher and participant should as far as possible be on equal terms. (Clearly, this consideration also holds for studying 'up' but in such cases the power to affect such equality often lies more with participants than with the researcher.) Thus, contracts between informed adult parties – often involving an interchange of material over time, for example views in an interview for views in a draft report – are seen as a more balanced exchange than the more traditional one-way subject-to-scientist transfer of data. It is also an exchange from which both parties can gain. As advocates, re-

(Continued on page 5)

(Continued from page 4)

searchers may gain from feedback offered by participants to the treatment of their material, and participants may gain from the sympathetic perspective of an 'outsider'.

Finally, much educational research is with children. Depending on their age, concepts such as contracts and informed consent may therefore apply differently. Children are often in a subordinate position to adult researchers, and adult carers may partly negotiate childrens' interests. In terms of comments on results, children may be of minimal use.

When are contracts possibly less appropriate?

Interestingly, some of the clearest examples of difficulties with contracts and member checks are in educational research. Perhaps because with children the power imbalance between researcher and researched is more obvious, some adults responsible for their welfare began in the 1970s negotiating de facto contracts.

Children are clearly less powerful than adults in the research process, and their interests are safeguarded by adults with the power in many cases to deny researchers access to the research setting as well as to take exception to findings. Exception is unlikely to be based on criticism of research participants. Educational research is usually constructive, rather than being intentionally destructive, of patriarchy or capitalism, for example. In any event, scientific critique is not *ad hominem*. Rather, institutions or structures may be critiqued. While much educational research is interested in improving existing knowledge and practices, it is possible for a researcher to find him or herself in a position where, if only by implication, a school, for example, or an educational model is critiqued. Here, the researcher may come up against the different standpoint of an adult supervisor of the project, and different interpretations may result. The rights of the researcher to evaluate and publish on terms acceptable to research institutions

such as universities may then be put under possibly legal pressure.

An early example of just such a position is Ball's (1981) *Beachside Comprehensive*, an ethnographic study of a comprehensive school in the UK. Ball's reflections on his study are related in Burgess (1984) and deal in part with difficulties encountered using an albeit tacit contract and an explicit program of respondent validation (an early term for member checks). It is possible that the looseness of the initial agreement between Ball and the headmaster of the school in which he was to undertake his research led to later problems with member checks. Permission was summarily granted without any clear understanding of the possibly critical nature of the research. As participants took issue with Ball's text during member checks, the headmaster commented that 'I don't think that many of the staff realize the critical nature of sociology as a discipline' (Ball 1984: 87). A formal contract might possibly have made this nature more clear.

This possibility raises the question of how much information is required to constitute informed consent? (I do not have space to discuss issues of informed consent involving children and parents.) Certainly, few standard contracts inform potential research participants that a critical stance could be taken at the reporting stage. Such disclosure could result in participation being declined or denial of access to a research setting. Is gaining informed consent worth this price? What of the long tradition of critical theory and research in social science?

Like Ball, I consider that greater effort needs to be made to ensure that possible research participants appreciate that there is no royal road to science and that critique is not pitched at the personal micro level but at the macro scale of general models. One way of introducing these concepts to research participants is to provide participants with a copy of a piece of the researcher's work, published if possible. This is obviously most appropriate with key

informants and with long or multiple interview participants, compared with perhaps hundreds of short questionnaire respondents. In the case of accessing school children, however, heads of schools and departments as well as teachers might all be informed in this way. Some possibly difficult differences of perspective might then be discussed at the beginning of the research process instead of emerging towards its end (clearly, changing research plans is likely to be easier at an early than a later stage in a research project).

Ball considers member checks might be fruitful if a free and open dialogue is possible between researcher and participants. He goes on to state, however, that researchers ignore at their peril 'the political aspects of dialogue ... Schools are political arenas where opposing ideologies and competing vested interests are played out. Any case study which taps into these facets of institutional life would seem to stand little chance of consensual agreement' (Ball 1984: 90). Conflict and compromise may be more prevalent in such cases than consensus, pointing to the issue of asymmetrical relations between a researcher and 'the powerful'. (There is insufficient space in this article to discuss different conceptualisations of power; it is important to note, however, that power can be both held and negotiated (e.g. Bradshaw 2001).)

The issue of the scale at which findings are written up in social science, touched on above, can also lead to misunderstanding between researcher and participants. Another way of conceptualising this issue is to distinguish between a work's internal and external audiences. It is common in social science to conceal the identities of individuals and even organisations, etc. participating in research. Ensuring confidentiality, however, like informed consent, may involve some discretion. Is there a point, for example, in changing the name in published work of a newsprint manufacturer in Australia when only one such enter-

(Continued on page 6)

(Continued from page 5)

prise exists? Just a small amount of knowledge or research on the subject would enable the identity of the enterprise to be deduced. In short, the more inside knowledge a reader possesses, the more difficult confidentiality can become. Participants are often in a position to attribute comments to particular individuals, despite the best efforts of a researcher to conceal these. In my experience, it did not matter to some participants when I denied their attributions; they had expected me to do so, they said, and remained convinced.

Conclusion

There is clearly a place for contracts between researchers and participants in social science. Explicating reciprocal rights and responsibilities at the start of research can head off later difficulties. Member checks can be mutually informative and empowering exercises that leave both researcher and participants better off for the experience. However, contracts, informed consent and member checks may apply differently where children are involved. In such instances, adult carers will most likely also be party to contracts. Researchers taking the position of constructive critic rather than advocate need to appreciate that member checks, in particular, risk leading to disagreement that can 'spoil the field' for all concerned.

Care should be taken therefore in assessing the appropriateness of contracts regarding a particular research setting, as well as to putting careful thought into any contract drawn up. As a preparatory document preceding research and underpinned by already developing relationships between researcher and key participants, contracts ought not to be approached lightly. Well-tailored contracts can smooth the research road, even building critical dialogue into the bargain through member checks. Operating beneath the general injunctions of ethics committees, judgement as to how open a researcher needs to be regarding a specific research setting is still largely the preserve of the researcher.

References

- Babbie E (1998) *The Practice of Social Research* (Wadsworth Publishing Company, Belmont), 8th edition
- Ball S (1981) *Beachside Comprehensive: A Case Study of Schooling* (Cambridge University Press, Cambridge)
- Ball S (1984) 'Beachside reconsidered: reflections on a methodological apprenticeship' in Burgess R (ed.) *The research process in educational settings: ten case studies* (The Falmer Press, London and New York), 69-96
- Baxter J and Eyles J (1997) 'Evaluating qualitative research in social geography: establishing 'rigour' in interview analysis' *Transactions of the Institute of British Geographers* 22, 505-525
- Berg B (1998) *Qualitative research methods for the social sciences* (Allyn

- and Bacon, Boston), 3rd edition
- Bradshaw M (2001) 'Contracts and member checks in qualitative research in human geography: reason for caution?' *Area* 33, 2, 202-211
- Burgess R (ed.) (1984) *The research process in educational settings: ten case studies* (The Falmer Press, London and New York)
- Cormode L and Hughes A (1999) 'The economic geographer as a situated researcher of elites' *Geoforum* 30, 299-300.
- Ellen R (ed.) (1984) *Ethnographic research: a guide to general conduct* (Academic Press, London)
- Guba E and Lincoln Y (1981) *Effective evaluation: improving the usefulness of evaluation results through responsive and naturalistic approaches* (Jossey-Bass Publishers, San Francisco)
- Guba E and Lincoln Y (1989) *Fourth generation evaluation* (Sage Publication, Newbury Park)
- Punch M (1986) *The politics and ethics of fieldwork* (Sage Publications, Beverly Hills)
- Witrock M (ed.) (1986) *Handbook of research on teaching* (MacMillan, New York and London), 3rd edition

What is an effect size?

Rob Coe

School of Education, University of Durham

Effect size is simply a way of quantifying the size of the difference between two groups. It is easy to calculate, readily understood and can be applied to any measured outcome in education. It is particularly valuable for quantifying the effectiveness of a particular intervention, relative to some comparison. It allows us to move beyond the simplistic, 'Does it work or not?' to the far more sophisticated, 'How well does it work in a range of contexts?'

Moreover, by placing the emphasis on the most important aspect of an intervention – the size of the effect – rather than its statistical significance (which conflates effect size and sample size), it promotes a more scientific approach to the accumulation of knowledge. For these reasons, effect size is an important tool in reporting and interpreting effectiveness.

The routine use of effect sizes, how-

ever, has generally been limited to meta-analysis – for combining and comparing estimates from different studies – and is all too rare in original reports of educational research (Keselman *et al.*, 1998). This is despite the fact that measures of effect size have been available for at least 60 years (Huberty, 2002), and the American Psychological Association has been officially encouraging authors to report effect sizes since

(Continued on page 7)

(Continued from page 6)

1994 – but with limited success (Wilkinson *et al.*, 1999). Formulae for the calculation of effect sizes do not appear in most statistics text books (other than those devoted to meta-analysis), are not featured in many statistics computer packages and seem to be seldom taught in standard research methods courses. For these reasons, even the researcher who is convinced by the wisdom of using measures of effect size, and is not afraid to confront the orthodoxy of conventional practice, may find that it is quite hard to know exactly how to do so.

Why do we need effect size?

Consider an experiment conducted by Dowson (2000) to investigate time of day effects on learning: do children learn better in the morning or afternoon? A group of 38 children were included in the experiment. Half were randomly allocated to listen to a story and answer questions about it (on tape) at 9am, the other half to hear exactly the same story and answer the same questions at 3pm. Their comprehension was measured by the number of questions answered correctly out of 20.

The average score was 15.2 for the morning group, 17.9 for the afternoon group: a difference of 2.7. But how big a difference is this? If the outcome were measured on a familiar scale, such as GCSE grades, interpreting the difference would not be a problem. If the average difference were, say, half a grade, most people would have a fair idea of the educational significance of the effect of reading a story at different times of day. However, in many experiments there is no familiar scale available on which to record the outcomes. The experimenter often has

Equation 1

$$\text{Effect size} = \frac{[\text{Mean of experimental group}] - [\text{Mean of control group}]}{\text{Standard Deviation}}$$

to invent a scale or to use (or adapt) an already existing one – but generally not one whose interpretation will be familiar to most people.

One way to get over this problem is to use the amount of variation in scores to contextualise the difference. If there were no overlap at all and every single person in the afternoon group had done better on the test than everyone in the morning group, then this would seem like a very substantial difference. On the other hand, if the spread of scores were large and the overlap much bigger than the difference between the groups, then the effect might seem less significant. Because we have an idea of the amount of variation found within a group, we can use this as a yardstick against which to compare the difference. This idea is quantified in the calculation of the effect size. The concept is illustrated in Figure 1, which shows two possible ways the difference might vary in relation to the overlap. If the difference were as in graph (a) it would be very significant; in graph (b), on the other hand, the difference might hardly be noticeable.

Calculating effect size

The effect size is just the standardised mean difference between the two groups (Equation 1).

If it is not obvious which of two groups is the 'experimental' (i.e. the one which was given the 'new' treatment being tested) and which the 'control' (the one given the 'standard' treatment – or no treatment – for comparison), the difference can still be calculated. In this

case, the 'effect size' simply measures the difference between them, so it is important in quoting the effect size to say which way round the calculation was done.

The standard deviation is a measure of the spread of a set of values. Here it refers to the standard deviation of the population from which the different treatment groups were taken. In practice, however, this is almost never known, so it must be estimated either from the standard deviation of the control group, or from a 'pooled' value from both groups. The issue of which standard deviation to use is discussed in more detail in Coe (2002).

In Dowson's time-of-day effects experiment, the standard deviation (SD) = 3.3, so the effect size was $(17.9 - 15.2)/3.3 = 0.8$.

Interpretations of effect size

One feature of an effect size is that it can be directly converted into statements about the overlap between the two samples in terms of a comparison of percentiles.

Interpretations of effect size

An effect size is exactly equivalent to a Z-score of a standard Normal distribution. For example, an effect size of 0.8 means that the score of the average person in the experimental group is 0.8 standard deviations above the average person in the control group, and hence exceeds the scores of 79% of the control group. With the two groups of 19 in the time-of-day effects experiment, the average person in the 'afternoon' group (i.e. the one who would have been ranked 10th in the group) might have scored about the same as the 4th highest person in the 'morning' group, since 79% (i.e. about 15) of the 19 are below this. Visualising these two individuals can give quite a graphic interpretation of the difference between the two effects. Table 1 shows conversions of effect sizes to percentiles.

Another implication of this interpretation

(Continued on page 8)

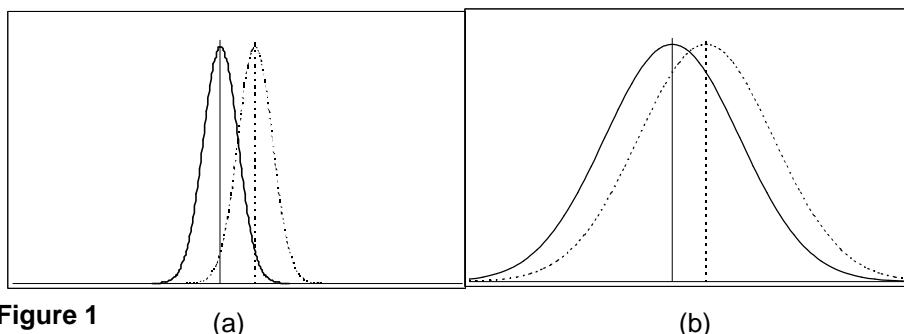


Figure 1

Table 1: Interpretations of effect sizes

Effect Size (standardised mean difference)	Percentage of control group who would be below average person in experimental group
0.0	50%
0.1	54%
0.2	58%
0.3	62%
0.4	66%
0.5	69%
0.6	73%
0.7	76%
0.8	79%
0.9	82%
1.0	84%
1.2	88%
1.4	92%
1.6	95%
1.8	96%
2.0	98%
2.5	99%
3.0	99.9%

(Continued from page 7)

tation is that changes in the percentage of people reaching a particular threshold can be translated to and from effect sizes. For example, if a school in which 50% of pupils were previously gaining five or more A* – C grades at GCSE were to implement a change with an effect size of 0.8, this percentage (other things being equal, and assuming that the effect applied equally across the whole curriculum) would rise to 79%. Even what would generally be thought of as a small effect of 0.2 would produce an increase from 50% to 58% – a difference that schools might well categorise as quite substantial.

A number of other interpretations of effect size have been put forward. These include conversions between effect sizes and correlation coefficients (Cohen, 1969), interpretations in terms of the difference in the percentage 'successful' in each of the two groups (the 'Binomial Effect Size Display', Rosenthal and Rubin,

1982) and the probability that a score sampled at random from one distribution will be greater than a score sampled from the other (the 'Common Language Effect Size', McGraw and Wong, 1992), among others.

It should be noted that the values in Table 1 depend heavily on the assumption of a Normal distribution. The interpretation of effect sizes in terms of percentiles is very sensitive to violations of this assumption. This issue is discussed in more detail in Coe (2002).

Conclusion

The above is intended to be a simple introduction to the idea of effect sizes. However, there are a number of issues surrounding their use, of which space does not permit discussion here. The reader may well have thought to question which standard deviation should be used in the calculation of an effect size, or what the consequences of calculating effect sizes which are based on samples with restricted ranges or which depend on unreliable measures may be. These questions are addressed in Coe (2002).

One important issue in justifying the dissemination of effect sizes is the use of confidence intervals around a calculated effect size to express the uncertainty in the estimate of a population value based on a single sample. In fact, confidence intervals generally convey the same information as the more widely used tests of statistical significance, but avoid the need for a usually inappropriate yes/no decision about whether there is an effect, instead allowing that effect to be quantified within a given margin of error. A further advantage of using effect sizes is that the results of different comparisons can be combined by a process of meta-analysis into an overall estimate of the size of the effect. This is a complex process, however, and, like most methods of analysis, can be done well or badly.

Overall, though, the more widespread use of effect size measures would represent an improvement in

the level of sophistication of the analysis and interpretation of quantitative data, and should certainly be encouraged. It is hoped that this article may help to do just that.

References

Coe, R. (2002) *It's the Effect Size, Stupid: What effect size is and why it is important*. Paper presented at the British Educational Research Association annual conference, Exeter, 12-14 September, 2002. [Available at Education-line, <http://brs.leeds.ac.uk/~beiwwww/beid.html>]

Cohen, J. (1969) *Statistical Power Analysis for the Behavioral Sciences*. NY: Academic Press

Dowson V. (2000) "Time of day effects in school-children's immediate and delayed recall of meaningful material". *TERSE Report* <http://www.cem.dur.ac.uk/ebeuk/research/terse/library.htm>

Huberty, C.J.. (2002) 'A history of effect size indices'. *Educational and Psychological Measurement*, 62, 2, 227-240

Keselman, H.J., Huberty, C.J., Lix, L.M., Olejnik, S. Cribbie, R.A., Donahue, B., Kowalchuk, R.K., Lowman, L.L., Petoskey, M.D., Keselman, J.C. and Levin, J.R. (1998) 'Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA analyses'. *Review of Educational Research*, 68, 3, 350-386

McGraw, K.O. and Wong, S.P. (1992) 'A Common Language Effect Size Statistic'. *Psychological Bulletin*, 111, 361-365

Rosenthal, R, and Rubin, D.B. (1982) 'A simple, general purpose display of magnitude of experimental effect.' *Journal of Educational Psychology*, 74, 166-169

Wilkinson, L. and Task Force on Statistical Inference, APA Board of Scientific Affairs (1999) 'Statistical Methods in Psychology Journals: Guidelines and Explanations'. *American Psychologist*, 54, 8, 594-604

The RCBN are holding a series of seminars on the role of numbers in teaching and learning research. The use of effect sizes can be explored further in these meetings. For more information see the back page and visit the RCBN website.

Handling occupational information

Paul Lambert

Cardiff University School of Social Sciences

Researchers regularly collect information on occupation, but are often apparently unaware how to utilise such information in an efficient and scientifically consistent way. In this note I discuss three practical issues in handling occupational information. Many of the comments make reference to internet resources listed in Table 1, which cover the kind of questions I am most frequently asked about (T.1 to T.7).

The preservation of two levels of data

At the heart of occupational analysis lies a tension between precise occupational details and derived occupational representations (such as social class categories or occupational gradings). At the point of data collection, information is usually recorded with a textual description of a specific activity. However by the point of analysis, it may be cognitively very difficult to fully incorporate such information, whereas some form of derived representation may be more easily understood, communicated and analysed. The most important argument of this note is that it is highly desirable, as well as relatively easy, to preserve and use both such levels of information.

We recommend researchers first save precise occupational details, then use recognised classification systems to derive subsequent representations, and below we sketch key elements relevant to achieving this in the UK. (Formal governmental advice on the measurement and coding of occupations for larger scale data is available via T.5). The two-level approach is desirable, first, because precise occupational details continue to play a role in many theoretical explanations. Second, because there is significant debate on how to represent occupations most appropriately, and the use of automatic coding frames linking precise units with one or more representations provides for a consistent and replicable derivation. In contrast, a common example of

'bad' practice in this field involves researchers hurriedly classifying their initially rich occupational data by hand, laboriously, into a single, simplistic classification system, often solely on the basis of their own personal judgements on occupational groupings. Subsequently, such data is analysed or released only at the level of this classification and often without documentation, rendering it ultimately non-comparable with other research.

Occupational units and source data

A fine level of occupational detail can be preserved by coding into 'occupational units', and the UK has for some time maintained 'official' numerically indexed unit classifications. These, latterly the 'Standard Occupational Classification' (SOC), have been regularly updated each decade, reflecting occupational restructuring; most recent for instance is the 'SOC2000' scheme, which incorporates 353 units. Documentation on these is found in successive HMSO volumes (such as ONS (2000) or T.1). A number of recent comments on SOC classifications can be found in Rose and O'Reilly (1997) and Rose and Pevalin (2002). The documentation usually features an alphabetical list of occupational titles – around 26,000 for SOC2000 – in a table format linked to the appropriate numerical SOC unit (see T.3). A classification of source data to SOC units can then be achieved manually from these look-up tables, and this may be an efficient procedure for users with a small number of cases, though software has also been developed to automate this process. The CASOC programme (see T.2) is designed to translate textual occupational descriptions into appropriate SOC (and other) occupational unit classifications. The first version (released 1993), uses a DOS prompt to categorise to SOC90 data, though it must be purchased for around £200. A revised CASOC, appropriate for both SOC90 and SOC2000, is planned to be completed around

December 2002, and will be freely accessible to small scale users via T.2, with input via a web browser interface.

Normally a coding to SOC units offers more than adequate occupational detail, and analysis can be undertaken immediately by identifying specific occupational units. Occasionally however, even finer differentiation is desired, most typically for reviews of specific occupations or populations. In one example relevant to educational research, 'HESA' (see T.6) publish a selection of refinements to SOC90 categories, known as "SOC(FDS)", which are designed to further differentiate selected occupations in a manner pertinent to recent leavers from higher education.

A rival occupational unit system is the International Standard Classification of Occupations (ISCO), also with several historical versions (see T.2, T.4). A motivation for classifying to ISCO is international comparability. Again, published ISCO volumes (e.g. ILO 1990) include look-up tables to code units. More generally, in many circumstances it is possible to translate both within and between different SOC and ISCO versions using coding frame computer files. Some such files can be supplied from the Occupational Information Unit (SOC90 or SOC2000 to ISCO88, see T.3), and more are available for access through links from the CAMSIS webpages (T.4). However in some cases the appropriate translations between different unit level classifications are not feasible due to basic substantive differences in their classifications. Most notably for the UK, strictly speaking there is no linear translation between the SOC90 and SOC2000 versions.

The SOC and ISCO unit classifications deal with occupational title information, but other types of occupational information are also often obtained. Firstly, occupations can

(Continued on page 10)

(Continued from page 9)

be classified by the industry to which they contribute. Standardised classifications of industries exist, in Britain named 'SIC80' and 'SIC92', cross-nationally harmonised versions including 'ISIC' and 'NACE'. T.5 includes a comment on the nature of the questions necessary to obtain industry information. An automated classification software of the same nature as CASOC is in preparation from the same authors (T.2), to classify textual descriptions into SIC80 and SIC92 units.

Secondly, occupational titles are often supplemented with several measures asked by direct questions. 'Employment status' is most widely premised on the division between employees and the self-employed, and T.5 lists the standard UK definition, which closely coincides with an international standard, 'ICSE'. Other measures often recorded include the size of the firm worked at, whether the individual has managerial or supervisory responsibilities, whether the individual is a trainee, whether they work full or part time, and whether on a temporary, fixed term or permanent contract (cf T.5). Information from such questions is often cross-classified with occupational unit information when specifying certain occupational representations.

Analytical representations: categorical classes and continuous scales

Much has been written on alternative possible occupational representations and it is difficult to summarise these fairly in a short note. Different approaches may be more or less appropriate to a particular theoretical explanation, so researchers should ideally consider their own approach in the context of literature on occupational stratification (see Crompton 1998 for a simple review). Below we mention five approaches which are widely recognised, easily operationalised from unit level data, and thus replicable.

Both occupational title and industrial unit classifications are hierarchically structured into successively fewer subgroups, so a linear translation of

either can immediately be used to group together different occupations into categories of 'sector' or 'industry' (some resources to achieve these subgroupings can be downloaded from T.4).

In the last few years, the pre-eminent UK occupational categorisation has become the NS-SEC (see T.1, and Rose & Pevalin 2002). This class scheme is founded on 'employment relations', primarily reflective of the also widely used 'Goldthorpe' or 'EGP' categorisations. Step-by-step instructions on producing NS-SEC categories from SOC-2000 unit group data (plus employment status information) are available from T.1. Production of earlier 'Goldthorpe' classifications can be achieved through files linked from T.4 (the files used for 'CAMCOM' for SOC90 data, or the link to Ganzeboom's pages for ISCO units).

Alternatively, measures ultimately related to the perceived 'skill' levels associated with jobs have, historically, been more widely used in Britain. Versions of the Registrar General's Social Class Schema may be derived from reclassifications of SOC units presented in the SOC codebooks (also incorporated into the 'CAMCOM' files from T.4). More recently Peter Elias has worked on producing skill level representations for simple recodes of the modern SOC and ISCO schemes, see links via T.2.

Lastly several approaches pertain to represent occupational structure in terms of a hierarchy, realised by assigning numerical scores to different occupational units. In the CAMSIS approach a hierarchy of relative stratification locations is defined on the basis of the patterns of social association between occupational unit holders. The webpages of this project, T.4, include introductions then extensive guidelines on ways of assigning appropriate scores to SOC and ISCO units. Alternatively, in work analysing the 'prestige' or 'socio-economic status' associated with occupations, Ganzeboom and Treiman released scores for derived representations of both concepts

(see links from T.4 [occunits/distribution.html]). An analytical advantage of these approaches is the production of a continuous variate representing occupational difference.

Households and gender

The concepts we try to represent through occupational information are interwoven with the demographic circumstances in which we live our lives. For instance the significance of an occupation to an individual may be contextual to the occupations of other household sharers. Also, it is well established that gender groups experience vastly different occupational distributions, which implies that the same occupation may have different implications for men and women. In this light, how occupational information is best analysed for individuals could be highly relative. However, a common analytical approach in Britain is unreflectively individualist and gender neutral: researchers derive equivalent representations regardless of gender, and largely ignore other household occupational structures. Indeed, for many this simplifying strategy appears to embody a political statement of opposition to the 'sexist' treatment allegedly inherent in any other method.

A long history of debates remain inconclusive over whether such an individualist approach is preferable or not (references can be found from Crompton's 1998 review). Firstly, regarding household context, a commonly presented alternative is to assign to individuals occupations held by themselves ('individualist'), or those held by the male household head when identified ('conventional'). Another is simply to incorporate multiple variables indicating both own occupations and those of household sharers. These alternatives to the individualist approach can be operationalised by simple file matching techniques utilising household identifiers. Secondly, many studies have presented evidence of extreme clustering and segregation in the occupational distributions of men and women, but in this case hardly any commonly used

(Continued on page 11)

(Continued from page 10)

occupational representations allow for separate male and female structures (one exception being the CAMSIS approach, T.4). At the very least, however, a basic consideration of demographic context can be made by separating gender groups during analysis (the same issues, furthermore, apply to other demographic differences characterised by similarly differential occupational distributions, for instance ethnic or regional groups).

Careers and current situations

Another relativistic element to the impact of occupations on individuals' lives involves the lifetime context of careers. At any given point of measurement, an individual may not presently hold the occupation which best summarises their general circumstances (this is particularly true for typical female careers). Again, a simplifying research strategy prevails in the UK which is not necessarily preferable. Most commonly, analysts consider the impact of occupations solely for currently employed individuals, in terms of their current occupations regardless of career stage.

More advanced treatments operate at both the levels of data collection and derived representations. Firstly, further unit level details can be obtained through data on the last main occupation of an individual, as is commonly obtained in many sur-

veys. If resources permit, collection of further descriptions of the career history of an individual provides better context still. Secondly, occupational representations can then be derived in a way which may incorporate lifetime positions more thoroughly. One possibility is to base an individual classification on either current, or last, main job, depending on whether currently working. Such a representation can then be analysed in the context of interaction effects with current employment activity information. Similarly, relative career location can be given a rudimentary consideration by adding the context of interaction effects with age or age cohort. More advanced methods, when career history information is available, have tried variously to represent 'average' career location; 'highest' lifetime career location; only the occupations of those in 'mature' employment, for instance aged over 35; or to classify career trajectories themselves (some relevant discussions can be found in the collection of Blossfeld & Drobnic 2001). There is no clear 'winning' prescription for considering career context in social science occupational analyses, whereas restricting analyses to the currently employed does at least ease the interpretation and communication of results. Moreover, the incorporation of career context is probably the most technically challenging element of occupational information discussed in this note. Nevertheless, the most appropriate

methodology is again likely to depend upon the theoretical questions asked.

Further issues

So far this note has referred to contemporary Britain, but it is also worth noting that similar issues apply to internationally or historically comparative research, where an added need for contextual awareness is introduced. A quick point is that most derived occupational representations are fixed in time and place and therefore should be contextualised by their 'marginal' distributions, although the CAMSIS representation does attempt to be deliberately relativistic in this regard.

Aside from looking further into the occupational literature, closing practical advice is to encourage researchers to follow previous data handling strategies, most obviously by reviewing the approaches of major UK datasets such as the census and government surveys, which are increasingly freely available and well documented (T.7).

References

- Blossfeld H-P, Drobnic S. (2001) *Careers of Couples in Contemporary Societies: From Male Breadwinner to Dual Earner Families*. Oxford: Oxford University Press
- Crompton R. (1998) *Class and Stratification: An Introduction to Current Debates, 2nd Edition*. Cambridge: Polity Press
- ILO. (1990) *ISCO-88: International Standard Classification of Occupations*. New York: International Labour Office
- ONS. (2000) *Standard Occupational Classification 2000, Volume 1: Structure and description of unit groups*. Office for National Statistics, London: HMSO
- Rose D, O'Reilly K, (eds.) (1997) *Constructing Classes: Towards a New Social Classification for the UK*. Colchester: University of Essex
- Rose D, Pevalin DJ. (2002) *A Researcher's Guide to the National Statistics Socio-economic Classification*. London: Sage

Table 1: Internet resources for handling occupational data

1. http://www.statistics.gov.uk/methods_quality/ns_sec/
Guidance on deriving NS-SEC classification from occupational sources
2. [http://www.warwick.ac.uk/ier/\[pubs/software.htm\]](http://www.warwick.ac.uk/ier/[pubs/software.htm])
Institute of Employment Research, includes links (via 'publications'; 'software') to news on CASOC, plus SOC90 descriptive files.
3. occupational.information@ons.gsi.gov.uk
Contact for the Occupational Information Unit, who can freely supply SOC2000 codebook file and other relevant SOC90/2000 information and files.
4. [http://www.cardiff.ac.uk/socsi/CAMSIS/\[occunits/distribution.html\]](http://www.cardiff.ac.uk/socsi/CAMSIS/[occunits/distribution.html])
CAMSIS project, includes translation files linking SOC units to CAMSIS scores, and, under 'occupational unit handling', a list of downloadable resources and links for processing occupational unit data for Britain and many other nations.
5. <http://www.statistics.gov.uk/harmony/>
Official recommendations on 'harmonised concepts and questions for government social surveys', including occupational information.
6. <http://www.hesa.ac.uk/>
The Higher Education Statistics Agency: Includes coding frame to SOC(FDS) via 'data collection' and 'coding manuals(FDS)'.
7. <http://www.data-archive.ac.uk/>
The UK Data Archive: access or browse existing datasets and their documentation

RCBN News

RCBN draft plan of activities now available

We have drawn up a provisional programme of capacity building activities that will currently run until May 2004. These capacity building activities have been structured around themes that capture the types of research capacity building most requested by members of TLRP and the wider educational research community. A diary of the events included in this programme is now available via the new RCBN activities website (see below).

RCBN sister-website launched

The RCBN website is updated regularly, and amongst the latest additions is a sister-website that will host information regarding the RCBN's own programme of activities. The site includes a diary of events, a list of activities by capacity-building theme, information on the learning networks that are being established for each capacity building theme (see below), as well as information regarding the latest resources available for RCBN activities.

Learning networks launched

The RCBN is setting up a learning network to support the capacity building activities of each of its themes. The first of these will be established at the New Career Researchers Conference in November 2002. The aim of the learning networks is to facilitate peer-led, virtual and face-to-face learning that outlives the RCBN. Membership will be open to all individuals who wish to pursue capacity building around a particular theme.

Visits to team leaders scheduled

The RCBN aims to consult widely with members of TLRP concerning the content and structure of its programme of activities. As part of this process, we are arranging to visit TLRP team leaders, to both consult on the provisional programme, and to encourage their par-

ticipation as methodological experts. We would also welcome your views of the proposed programme of activities.

Support for Phase III

The RCBN provided methodological assistance to research teams shortlisted for TLRP Phase III. This support will be continued by including Phase III projects in the RCBN's capacity-building activities.

Forthcoming RCBN activities

The role of numbers

The RCBN have designed a 2-hour seminar that aims to encourage researchers to use numerical-datasets, to be less fearful of such data, and to overcome disruptive methodological identities. We are offering to hold this seminar at all institutions involved in TLRP. Seminar dates are currently set for:

University of Leeds, 28 November 2002
Middlesex University, 22 January 2003
University of Durham, 29 January 2003
University of Wales Institute, Cardiff, 4 February 2003
University of Leicester, 6 February 2003
Staffordshire University, 5 March
University of the West of England, 14 March

These seminars are open to all members of TLRP, the host and neighbouring institutions, and policy makers and practitioners.

For more information on these seminars and an update on other seminars please visit our website. If you wish to find out which seminar is nearest to you, then please contact the Cardiff office.

ESRC Teaching and Learning Research Programme Research Capacity Building Network

Project Director

Stephen Gorard
gorard@cardiff.ac.uk

Project Administrator

Helen Taylor
taylorh1@cardiff.ac.uk

Building Research Capacity

ISSN 1475-4193, Issue 4

Editors

Stephen Gorard, Project Director
Chris Taylor (TaylorCM@cardiff.ac.uk)
Karen Roberts (RobertsK@cardiff.ac.uk)

Correspondence

Cardiff University School of Social Sciences
Glamorgan Building
King Edward VII Avenue
Cardiff CF10 3WT
Tel. +44(0)29 2087 5345
Fax. +44(0)29 2087 4678

Cardiff University School of Social Sciences

Views expressed in this journal do not necessarily reflect those of the University, the Research Capacity Building Network, the Teaching and Learning Research Programme, nor the Economic and Social Research Council.