

Public perceptions of unreliability in examination results in England: a new perspective

Qingping He¹ & Dennis Opposs

(Office of Qualifications and Examinations Regulation, Coventry CV5 6UB, UK)

Andrew Boyle

(The City and Guilds of London Institute, London EC1A 9DD, UK)

Abstract

The Office of Qualifications and Examinations Regulation (Ofqual) in England is carrying out a two-year research programme investigating the reliability of results from national tests and public examinations. One strand of the programme is to gauge public perceptions of unreliability in examination results. Based on findings from previous qualitative studies involving the use of workshops and focus groups, a further quantitative study on public perceptions of reliability using an online questionnaire survey was conducted. The questionnaire was structured into five distinctive topics to measure different aspects of respondents' knowledge of and attitudes towards unreliability in examination results. Respondents were sampled from three key stakeholder groups: A level teachers, A level students aged 16-18, and employers. Substantial variability exists in the understanding of reliability concepts and attitudes to unreliability among the respondents. The level of tolerance of the respondents for measurement uncertainty to some degree was correlated to the level of belief about the examinations system, knowledge of aspects of unreliability and approaches to trust.

Keywords: Public perceptions, reliability, measurement error.

Background

Reliability, in educational measurement terms, refers to the consistency of results on a given measure from repeated measurements under equivalent conditions and is an important indicator of the quality of an assessment. However, there has been little large-scale research to monitor the reliability of results from England's test and examination systems and little understanding of the public's knowledge of and attitudes towards unreliability in assessment results. To address this, Ofqual is conducting a two-year research programme. The primary aim of this programme is to gather evidence to inform Ofqual on developing regulatory policy on reliability. The programme is structured into three strands:

- Strand one: Generating evidence of reliability;
- Strand two: Interpreting and communicating evidence of reliability;
- Strand three: Exploring public perceptions of reliability and developing policy on reliability.

It is essential to understand the public's attitudes towards uncertainty in examination results when developing regulatory policy.

Existing studies on public perceptions of reliability in England

As part of Strand three of the reliability programme, Ofqual commissioned Ipsos MORI and the Assessment and Qualifications Alliance (AQA) to carry out two qualitative research projects. These investigated public understanding of reliability and their opinions about the national examinations system and measurement error in examination results. The research focused on different aspects of reliability, including the assessment process; factors affecting the performances of students on exams; the reliability concepts and measurement error; the different types of error in examination results – preventable human mistakes versus inevitable random measurement error; factors contributing to measurement error in examination results; and the level of acceptance towards human error and measurement error in examination results.

Ipsos MORI (2009) used two workshops in its investigation. The workshop sessions started with an analogy to an error occurring in medical treatment; this constituted a substantial input to help workshop participants understand the concepts under discussion. The research participants appeared to accept that a certain amount of error was inevitable in a large examinations system, but they could be intolerant of “preventable errors” (Ipsos MORI, 2009). At times, participants appeared to be making a

¹ Email: Qingping.He@Ofqual.Gov.Uk

distinction between inherent and preventable error. Some research participants stated that their attitude to error depended on whether the error changed a student's grade or mark. There was evidence that students were aware that some inconsistency between human markers was inherent in subjects, such as English. However, there were also statements that such inherent error should be minimised or even eliminated.

Chamberlain (2010) from AQA conducted qualitative research to follow up Ipsos MORI's (2009) work. Like Ipsos MORI, Chamberlain designed her research with the assumption that she would have to take steps to mitigate participants' lack of knowledge of key elements of the reliability concept. Chamberlain used vignettes as a technique to introduce reliability to her research participants. The participants tended to be fairly trusting of the examinations process; trusting in the professionalism and training of subject experts. Participants felt it would be useful for reliability information to be communicated to the public in general terms, but were opposed to specific quantification of unreliability (e.g. via an indication of the amount of uncertainty associated with a grade) on a candidate's examination certificate. Chamberlain (2010) also suggested a series of hypotheses that could be addressed in a subsequent quantitative questionnaire survey.

Ipsos MORI conducts a survey of perceptions of A levels and GCSEs each year for Ofqual that is now in its eighth wave (Ipsos MORI, 2010). Findings from the 2010 survey suggested that the majority of teachers thought that most students got the correct grade at GCSE. However, the general public were more sceptical, with more respondents believing that a larger proportion of candidates got the wrong grade. Respondents' perceived reasons for candidates to get the wrong grade in examinations included: students performing better or worse than expected in examinations or coursework, inaccurate marking and poorly designed examination papers.

Studies on trust

Both the Ipsos MORI and the AQA qualitative work (Ipsos MORI, 2009; Chamberlain, 2010) suggested that some participants had limited awareness of reliability concepts. It would be likely that some of the research participants might have different attitudes to unreliability if they had not gone through the workshop/focus group process. It is likely that when one does not have full knowledge about a situation, attitude would be influenced by factors such as personal experience and approaches to trust and others. A brief literature review on studies of trust in various areas was also conducted to provide some insight into factors that affect trust, which would also apply to research into public trust in the reliability of examination results (see Putnam, 1995; Fukuyama, 1995; Hardin, 2002; O'Neill, 2002; Tschannen-Moran & Hoy, 2000; McLeod, 2002; Lycan, 1999; Coren et al., 1999; Reiss, 2000; Wilmot et al, 2005; Bradberry, 2007; Wilkinson & Pickett, 2009).

The present study

The analogy and vignettes used by Ipsos MORI (2009) and Chamberlain (2010) in their studies might have helped the participants to understand the concept of reliability and the factors that could introduce uncertainty in examination scores and to develop views on measurement error. The group discussions could also have influenced the opinions of the participants about error in examination results. Furthermore, the small sample size of these studies makes it inappropriate to make any generalisation of the findings. The Ipsos MORI (2010) survey only addressed some narrow aspects of reliability of examination results. The present study seeks to contribute further to a developing understanding of attitudes to reliability and unreliability using an objective questionnaire survey. It explores the public's awareness of and opinions about reliability in the following areas:

- Knowledge of and experience in the examination process and confidence in the national examinations system;
- Understanding of factors that affect the performances of students on examinations and factors that introduce uncertainty into examination scores;
- Attitudes towards different types of assessment error (including human mistakes and measurement inaccuracy);
- Approaches to trust in general.

Data collected is also used to investigate:

- How attitudes to unreliability are related to knowledge and understanding of the reliability concept;

- How attitudes to unreliability are related to confidence and belief in the examination system and approaches to trust;
- How confidence and belief in the examination system are related to trust.

Methodology

Instrument development

The questionnaire was structured into five distinctive topics:

- Topic A: Experiences of, and knowledge and beliefs about, the examinations system
- Topic B: Awareness of unreliability
- Topic C: Attitudes to unreliability
- Topic D: Views on approaches for improving reliability
- Topic E: Approaches to trust

This structure of the questionnaire allowed the balance in items between the topics to be controlled and for relationships between topics to be investigated. The questionnaire had 23 questions, many of which were multi-part, making a total of 80 individual sub-items. All sub-items were multiple choice questions (MCQs) with varying response options. Most sub-items required respondents to endorse their views on a statement with the responses varying from “Strongly agree” to “Strongly disagree”. A few questions also used “Don’t know” as a response category. The statements were varied, to contain positive and negative statements (Pearson & Raeke, 2000).

Respondents sampling

Three groups of stakeholders were chosen to provide samples of respondents to the questionnaire: sixth form students studying on A level courses in schools or colleges in England; school teachers who teach on A level courses in schools or colleges in England; and employers (especially members of staff with responsibility for recruitment).

The National Foundation for Educational Research (NFER) was contracted to sample respondents and administer the questionnaire to collect data. To obtain responses from teachers and students, samples of institutions were drawn from the NFER’s Register of Schools and Colleges. A random sample of 800 institutions was drawn from maintained and independent schools in England with a representative number of sixth form centres and Further Education (FE) colleges. The 800 institutions in the sample included only those containing year 12 and year 13 students. Two or three A level teachers and five or six A level students were invited to complete the online questionnaire. Experian provided the sample of employers. It was anticipated that it could be difficult to engage employers in research, so representatives were invited from 3,000 companies. The achieved sample sizes were 314 for teachers, 358 for students and 210 for employers, resulting in the error associated with the population estimate of the percentage response to a sub-item to be estimated at about $\pm 5.5\%$, $\pm 5.2\%$ and $\pm 6.8\%$, respectively, at a 95% confidence interval.

Data collection and analysis

The questionnaire was administered via the internet. All data were collected anonymously so that no organisations or individuals could be identified in subsequent analysis. To facilitate statistical analysis, the response categories in a sub-item were transformed into numerical values, varying from 0 for the weakest category, to the number of options minus 1 in the sub-item for the strongest category. For negatively asked question, the transformation was reversed. The coded data were analysed for reliability for each topic for each of the respondent groups, in addition to analysis for some basic descriptive statistics at both sub-item level and topic level. Correlation analysis between the topics for each group was also conducted to investigate how attitudes to assessment error correlate with other attributes of the respondents based on information obtained from individual topics.

Results and Discussion

Instrument internal consistency reliability

The questionnaire used in the present study is a multi-dimensional instrument; it attempts to explore respondents’ approaches to trust, knowledge and understanding of the assessment process; factors affecting students’ performances on exams; factors introducing errors in examination scores; and

attitudes to unreliability in examination results. The internal consistency reliabilities of scores represented by Cronbach’s alpha for the individual topics for each group were reasonably adequate (with values varying from 0.58 to 0.85), except for Topic D (Views on approaches for improving reliability) for teachers (a value of 0.25 for alpha), suggesting that there was greater variation in choosing the response categories of sub-items in Topic D for teachers. Topic D was also the shortest topic in the questionnaire, containing 9 sub-items.

Confidence in the national examinations system

Questions in the first topic concerned respondents’ personal experience, knowledge and opinions about the national examinations system. There were 23 sub-items in this topic.

For the statements “In general, students get the grades they deserve in exams” and “To what extent do you agree or disagree with the statement ‘I have confidence in the national examination system’?”, about 89% of the teachers felt that their students got the grades they deserved, and over 66% of the employers thought that students got the grades they deserved. These findings are broadly in line with the findings from the 2010 Ipsos MORI survey and the qualitative studies by Ipsos MORI (2009) and Chamberlain (2010). About 62% of the teachers showed confidence in the examinations system, which is lower than the percentages from the 2010 Ipsos MORI survey about views on the accuracy of GCSE grades. The percentages of students and employers who had confidence in the system were substantially lower than those of teachers, at 42% and 39%, respectively. Teachers are likely to be more confident than students and employers are in the examinations system, as they use the system more and are more familiar with it.

When asked about their opinions about the performance of the national exams system, 26% of teachers, 25% of students and 18% of employers felt that the system was doing either a very good job or a good job. However, about 61% of teachers, 57% of students and 48% of employers thought that the system was doing a good job but needed improving. Again, teachers trusted the system more than students and employers. About 12% of the teachers, 14% of the students and 23% of the employers thought that the system was not doing a good job and should be reformed.

Table 1 shows percentages of respondents from the three groups who regarded the various factors as important for creating trust in the examinations system. The vast majority of respondents thought all the four factors listed in Table 1 important in creating trust. The endorsement rate for “Subject experts making sure that examinations measure the right things and they are at the right level” and “Examination boards have the necessary expertise and experience” was over 91%. In terms of use of statistical procedures in awarding, the endorsement rate was over 80% for teachers and students and about 74% for employers. These findings were generally consistent with the findings from Chamberlain’s work, which suggested that the participants involved in the focus group discussions trusted the examinations system because they believed that awarding bodies had procedures in place to ensure that candidates received the grades they deserved.

Table 1 Percentages of respondents endorsing factors that create trust in the examinations system.

Factors	Teachers	Students	Employers
Subject experts making sure that examinations measure the right things and that they are at the right level	96%	94%	93%
Statistical procedures making sure that examinations give the right results	81%	88%	74%
The examinations system is a national system	91%	85%	86%
Examination boards have the necessary expertise and experience	94%	92%	92%

Understanding of factors affecting performances on examinations and factors causing unreliability in examination results

Questions in the second topic of the questionnaire concerned respondents’ awareness of reliability issues, including questions about understanding those factors that can affect students’ performances on examinations and/or introduce errors in examination scores. There were 12 sub-items in this topic.

The studies conducted by Ipsos MORI (2009) and Chamberlain (2010) indicated that the research participants generally understood the many factors that could affect the performances of students on exams. This was further confirmed by findings from the present study. In general, all five factors listed in the questionnaire were regarded as important in influencing students’ performances on exams. Of

these factors, knowledge about the subject and preparedness of the student were regarded as most important by all respondents (with endorsement rates over 91%). “How well the student feels on the day” and “Who marks the question paper?” were regarded as less important than knowledge about the subject and preparedness, with the endorsement rate varying from 57% for employers to over 66% for teachers and students.

The study by Ipsos MORI (2009) suggested that students were aware that some inconsistency between human markers was inherent in some subjects, including English. However, there were also statements that such inherent error should be minimised or even eliminated. Although the workshops were guided, some participants were still not quite clear about the factors that could introduce inconsistency in examination scores if the examination procedure was repeated. Results from the present investigation seemed to indicate that the majority of the respondents, to some degree, understood the main sources of error in examination scores. Over 58% of the respondents selected either “Strongly agree” or “Agree” for the statements about factors that could cause inconsistency in examination results if the examination procedure was repeated. About 75% of the teachers and 85% of the students thought “Test questions (e.g. if a different test had been set, the student might not have been disadvantaged by the wording of an essay question)” an important error-contributing factor. Over 71% of teachers and students felt that “Marking inconsistency (e.g. if a different marker had been assigned, the student might have achieved a different result)” could introduce error in examination results.

Attitudes towards unreliability and assessment error

Questions in Topic 3 asked about respondents’ attitudes towards unreliability in examination results, including their tolerance for human mistakes and inevitable measurement uncertainties. There were 16 sub-items in this topic.

For the two statements about “error in examination grades and inaccuracy in the assessment system” and “inevitable inherent variability in examination results and avoidable human mistakes in the examination system”, over 64% of teachers and students selected “Examination results are essentially an estimate – a certain amount of error is inevitable” on one hand, but 56% of the teachers and 51% of the students also selected “All inaccuracy has to be removed from the system, there’s no such thing as ‘inevitable and acceptable variation’”, suggesting intolerance for error. This inconsistency may reflect the weak relationship between knowledge about reliability and attitudes to unreliability and is consistent with findings from the Ipsos MORI research. Employers, unlike teachers and students, were more intolerant of error.

Ipsos MORI’s research (2009) indicated that some participants’ attitudes to error depended on whether the error changed a student’s grade or mark. They considered grade-related error to be more consequential than mark-related. These findings are supported by findings from the present study (see Table 2 for endorsement rates for statements about attitudes to measurement error). Less than 49% of the respondents from all the three groups agreed that “Error in the *mark* a student receives which does not affect a grade overall is not a cause for concern”, while over 90% agreed that “Error which results in a student receiving a different *grade* to the one they deserve is serious”. Over 86% of the respondents from the different groups felt that “Error that changes a grade C to a grade D in a GCSE examination is particularly important”.

Table 2 Percentages of respondents endorsing statements about attitudes to measurement error.

Factors	Teachers	Students	Employers
Error in the <i>mark</i> a student receives which does not affect a grade overall is not a cause for concern	39%	37%	48%
Error which results in a student receiving a different <i>grade</i> to the one they deserve is serious	98%	91%	94%
Error that changes a grade C to a grade D in a GCSE examination is particularly important	94%	87%	87%

Over 76% of the teachers and employers endorsed the statements “The performance on the day of an examination can be affected by feeling stressed or unwell, but it is ‘just one of those things’”, “Students need to be held accountable for how they perform on the day of the exam” and “Examination boards should do everything they can to minimise inconsistency from their processes”. Unsurprisingly, only 53% of the students agreed that “Students need to be held accountable for how they perform on the day of the exam”, while 78% of the teachers and 77% of the employers endorsed

the statement. Over 94% of all the respondents agreed that “Examination boards should do everything they can to minimise inconsistency from their processes”. These findings again are generally consistent with the findings from the studies by Ipsos MORI (2009) and Chamberlain (2010).

Approaches for improving reliability

Topic 4 had 9 sub-items concerning respondents’ opinions about approaches that can be adopted to improve the reliability of examination results. Over 92% of the respondents agreed that “Improve training for markers” was important. Also, over 79% of all respondents agreed that “Have two markers for essays” was important. Interestingly, about 33% of the teachers felt it necessary to “Use more teacher assessment for awarding qualifications” to improve reliability, which rose to over 53% for employers and over 64% for students. This could suggest that teachers did not have great confidence in teacher assessments, while students and employers did. Only 18% of the teachers and 31% of the employers agreed to “Use more multiple-choice questions” to improve reliability, while the endorsement rate for students was about 50%. About 22% of the teachers, 17% of the students and 24% of the employers agreed to “Have longer tests” to improve reliability.

Findings from Chamberlain’s work (Chamberlain, 2010) indicated that the majority of the participants at the focus groups did not favour reporting of reliability statistics and believed that doing so would undermine candidates’ achievements and create uncertainty, although some participants suggested that the public should be informed of errors in examination results. Secondary school teachers felt that teachers and students needed to be better informed. These findings were supported by results from the present investigation. To explore this further, one question in Topic 4 asked for respondents’ views on whether uncertainties associated with examination grades should be indicated on a student’s certificate. About 67% of teachers, 33% of students and 52% of employers thought error associated with a grade should not be indicated on the certificate.

Approaches to trust

Questions in Topic 5 concerned respondents’ approaches to trust. There were 20 sub-items in this topic. Over 88% of the respondents from all the three groups selected either “Agree strongly” or “Agree somewhat” to the statements “I trust organisations if I have personal experience of them” and “I trust professionals with whom I come into personal contact”. Over 58% of the respondents selected “I trust organisations that have a strong technical focus”.

Relationships between belief, knowledge and approaches to trust and attitude to unreliability

The data collected were also used to explore the relationship between respondents’ attitudes to unreliability in examination results and their attributes like belief about the examinations system, knowledge about reliability concepts and approaches to trust. The internal consistency reliability of a topic, represented by Cronbach’s alpha, to a certain degree reflects the unidimensionality of the topic in measuring the underlying construct, and values of Cronbach’s alpha suggest that all the topics for the three groups had reasonably adequate internal reliabilities except for Topic D (Views on approaches for improving reliability) for teachers. An attempt was made to investigate the relationships between the topics, and Tables 3–5 list the correlation coefficients between topic scores for the three groups. Significant correlation exists between Topic C (Attitudes to unreliability) and the other topics, indicating the influence of knowledge and beliefs and approaches to trust on attitudes to unreliability. The magnitudes of the correlations reflected the degree of the effect of the various attributes on attitudes to reliability. In view of the relatively low level of reliability of the topics, the values of unattenuated correlation coefficients for the individual topics would be substantially higher than those listed in Tables 3–5.

Table 3 Correlations between scores on different topics for teachers

	Topic A	Topic B	Topic C	Topic D	Topic E
Topic A	1				
Topic B	0.227**	1			
Topic C	0.063	0.152**	1		
Topic D	0.056	0.153**	0.143*	1	
Topic E	0.192**	0.016	0.111	-0.017	1

*significant at $p < 0.05$, ** significant at $p < 0.01$

Table 4 Correlations between scores on different topics for students

	Topic A	Topic B	Topic C	Topic D	Topic E
Topic A	1				
Topic B	0.352**	1			
Topic C	0.125*	0.219**	1		
Topic D	0.222**	0.317**	0.337**	1	
Topic E	0.271**	0.296**	0.296**	0.328**	1

*significant at $p < 0.05$, ** significant at $p < 0.01$

Table 5 Correlations between scores on different topics for employers

	Topic A	Topic B	Topic C	Topic D	Topic E
Topic A	1				
Topic B	0.433**	1			
Topic C	0.378**	0.406**	1		
Topic D	0.288**	0.378**	0.341**	1	
Topic E	0.368**	0.233**	0.259**	0.194**	1

*significant at $p < 0.05$, ** significant at $p < 0.01$

Concluding Remarks

There has been little large-scale research to monitor the reliability of results from national tests and public examinations in England and limited understanding of the public's knowledge of and attitudes towards unreliability in examination results. The Ofqual Reliability Programme was designed to address these issues, which is important for improving the quality of the examinations system further. It is essential to understand the public's attitudes towards uncertainty in examination results when developing regulatory policy on reliability in order to increase their confidence in the examinations system. Results from this study indicated that knowledge about and attitudes to unreliability in examination results vary between respondents for the three stakeholder groups investigated. Most respondents from the three groups appeared to understand the assessment process and the factors that affect students' performances on exams. The respondents, to a degree, also understood the factors that could introduce uncertainty in examination results. The respondents showed various degrees of acceptance of measurement error in examination results.

Results from this study indicated that respondents' attitudes to unreliability positively correlated to their knowledge about aspects of reliability, beliefs about the examinations system and approaches to trust. A substantial proportion of respondents from the three groups lacked awareness of some aspects of reliability. This was also recognised by many assessment experts (see Ofqual, 2009; see also Boyle et al., 2009). Further study in this area would involve experiments to investigate how attitudes to unreliability could be affected by greater understanding of aspects of reliability. It is also important to explore effective ways of educating the public to understand reliability concepts and about uncertainty in examination results.

Although the findings from this study generally supported the findings from the qualitative investigations by Ipsos MORI (2009) and Chamberlain (2010), the differences in the findings from the two approaches have to be recognised. While the views expressed at the workshops or focus group discussions were under a controlled environment, the self-reported attitudes through the responses to a questionnaire excluded external influences. It is very likely that the use of workshops or focus groups would have helped the participant to develop knowledge and views about reliability. However, since the attitudes of the participants were not measured before and after the workshops/focus group process, it was impossible to assess the impact of the increase in knowledge about reliability on the change in their attitudes. As noted previously, the present study, to a certain degree, established the relationship between attitudes to unreliability and knowledge and other attributes of the respondents.

The present study was restricted to only three groups: teachers, students and employers, to whom the reliability of examination results would probably be more important than to other groups. Further

research would involve studying the perceptions of reliability from other stakeholder groups such as parents and the general public.

Acknowledgement: The authors gratefully acknowledge the assistance from Jo Taylor, Laura Daly and Jennifer Jupp for project management and data collection.

References

- Boyle, A., Opposs, D. & Kinsella, A. (2009) No news is good news? Talking to the public about the reliability of assessment. Paper presented at the 35th *International Association for Educational Assessment (IAEA) Annual Conference* in Brisbane, Australia, 13–18 September, 2009. Available online: <http://www.ofqual.gov.uk/files/2009-09-iaea-no-news-is-good-news.pdf>
- Bradberry, T. (2007) *The personality code*. New York: Putnam.
- Chamberlain, S. (2010) *Public perceptions of reliability*. Available online: http://www.ofqual.gov.uk/files/Ofqual_10_4708_public_perceptions_reliability_report_08_03_10.pdf
- Coren, S., Ward, L.M. & Enns, J.T. (1999) *Sensation and Perception*. New York, NY: Harcourt Brace.
- Fukuyama, F. (1995) *Trust: social virtues and the creation of prosperity*. New York, NY: Free Press.
- Hardin, R. (2002) *Trust and trustworthiness*. New York, NY: Russell Sage Foundation.
- Ipsos MORI (2009) *Public perceptions of reliability in examinations*. Available online at: http://www.ofqual.gov.uk/files/2009-05-14_public_perceptions_of_reliability.pdf
- Ipsos MORI (2010) *Perceptions of A levels and GCSEs □ Wave 8*. Available online: www.ofqual.gov.uk/files/2009-02-26-ofqual-perceptions-of-alevels-gcses.pdf
- Lycan, W.G. (ed.) (1999) *Mind and cognition: an anthology*, 2nd Edition. Malden, Mass: Blackwell Publishers.
- McLeod, C. (2002) *Self-trust and reproductive autonomy*. Cambridge, MA: MIT Press.
- Ofqual (2009) *The Reliability Programme – Technical seminar report*. Ofqual: Coventry, UK. Available online: <http://www.ofqual.gov.uk/research-and-statistics/research-reports/92-articles/20-reliability>
- O'Neill, O. (2002) *A Question of Trust: the BBC Reith Lectures 2002*. Cambridge: Cambridge University Press.
- Pearson, S.D. & Raeke, L.H. (2000) Patients' trust in physicians: many theories, few measures, and little data. *Journal of General Internal Medicine*, 15(7), 509–513.
- Putnam, R.D. (1995) Bowling alone: America's declining social capital, *Journal of Democracy*, 6(1), 65-78.
- Reiss, S. (2000) *Who am I?: the 16 basic desires that motivate our actions and define our personalities*. New York: Tarcher/Putnam.
- Tschannen-Moran, M., & Hoy, W.K. (2000) A multidisciplinary analysis of the nature, meaning, and measurement of trust, *Review of Educational Research*, 70, 547-593.
- Wilkinson, R. & Pickett, K. (2009) *The spirit level*. London: Allen Lane.
- Wilmot, A., Jones, J., Dewar, A., Betts, P., Harper, R. & Simmons, E. (2005) *A qualitative study on behalf of the Office for National Statistics and the Statistics Commission*. Available online: <http://www.ons.gov.uk/about/who-we-are/our-services/data-collection-methodology/reports-and-publications>