

Writing by numbers: are readability formulas to clarity what karaoke is to song?

Martin Cutts, Plain Language Commission, UK, mail@clearest.co.uk and <http://www.clearest.co.uk>. Edition 4, February 2008.

Based on a paper given by the author to the Plain Language Association InterNational (PLAIN) conference, Amsterdam 2007.

Summary

Readability tests such as SMOG (Simplified Measure Of Gobbledygook) offer an easy and relatively cheap way of assessing the apparent level of difficulty of a document. Their simplicity and veneer of mathematical certainty make them attractive propaganda tools. But bodies using the tests for propaganda, like the UK Government's education department, should mention their drawbacks when doing so, because the tests are crude. Using SMOG, one 10-sentence passage with (say) 12 polysyllabic words will score virtually the same as any other with 12 polysyllabic words. This means that the test ignores most of the criteria by which the clarity of documents is usually judged, such as use of appropriate headings, logical argument, lack of ambiguity, short sentences, good grammar and good punctuation. Assessing a document's clarity on test results alone is therefore likely to be misleading. It is not good if public policy is being shaped by such results nor if the news media are persuaded to give them undue credence.

This paper says discussion groups and one-to-one interviews are better than readability testing for assessing clarity. However, these methods are usually feasible only where cost and speed are not of great importance. This means that in deciding whether a document is at an appropriate level for its audience, editorial judgment based on experience will be the best aid. That judgment may, of course, take account of evidence such as a readability-test score.

From the confusion of evidence about adult reading abilities in the UK, the paper suggests that the average reading age may be 12–14 years – say equivalent to an average 13-year-old with adult experience. At 3 years below the school-leaving age, this is not very high but it is far higher than the average of 9.5 years quoted by some. An average of 9.5 for an adult is below functional literacy, if the National Literacy Trust's figures (appendix B) are correct. It seems self-evident that the average adult in the UK is not illiterate.

Introduction

The merits of readability formulas (also called readability tests) are hotly debated. Some people regard them as the essential starting point in any quest for clarity, others as a mathematical hoax, and others as an encouragement to bad writing because to get a 'good' test score authors need to reduce their syllable count and shorten their sentences, and they may take this too far. This paper looks at:

- 1 dubious uses of readability tests
- 2 what readability tests won't tell you
- 3 whether discussion groups are more useful than readability tests
- 4 the problem of irreducibility in plain-language work
- 5 the notion of 'average reading age' (UK) and 'average grade level' (US)
- 6 whether plain-language principles are relevant to works of literature
- 7 a new resource Plain Language Commission is developing for authors and editors, based on the *Living Word Vocabulary* that underpins several well-known readability tests.

1 Dubious uses of readability tests

At Christmas 2006, the UK Government's Department for Education and Skills (DfES) voiced its annual worry about low literacy standards. It said that millions of Britons would be unable to sing along to their favourite karaoke tunes during the new-year celebrations because weak literacy skills left them puzzled by the words on the screen. Its press release was headlined 'Millions facing lyrical letdown this Christmas'; and 'England's adults lack the skills to belt out karaoke favourites'. These assertions were uncritically reported by the news media, though the *Education Journal* (issue 100, 2007-01, page 10) called the whole thing 'rubbish', 'a gimmick' and 'nonsense'. The DfES claims were based on a SMOG readability test of 10 pop-songs and it concluded – from the test results alone – that 17.8 million adult Britons (out of around 39 million) would therefore be unable to read the lyrics of Robbie Williams' song 'Angels' as they came on the screen.

Part of the absurdity of this claim is that karaoke is a pastime of only a tiny minority of, mainly, pub-goers. As they have often partaken of strong drink pre-performance, getting the words muddled is not unknown – indeed, it's almost expected. Yet no incidents have been reported of singers being booed off stage for illiteracy. (In 2007, however, the doorman of a Philippines night-club did shoot to death a singer who was persistently out of tune.)

The SMOG test is based on the number of syllables in a 10-sentence sample. See <http://www.literacytrust.org.uk/campaign/SMOG.html>. For this paper I have put various lyrics (appendix A) through readability tests. The figures that follow are all US grade level – add 5 for the UK reading age – as scored using Micro Power & Light Co's Readability Calculations Version 6.0. The first 2 lyrics listed were also tested by the DfES.

- 'Angels' comes out at 4.1 on the Flesch–Kincaid test, but 7.5 on SMOG.
- 'Dancing Queen' by Abba scores 2.9 on Flesch–Kincaid and 6.8 on SMOG.
- 'Three Times a Lady' by Lionel Richie scores 2.2 on Flesch–Kincaid and 6.5 on SMOG.
- 'White Room' (rather less of a karaoke favourite) by Pete Brown and Jack Bruce

scores 3.7 on Flesch–Kincaid and 6.2 on SMOG.

It's interesting that 'Three Times a Lady', which has incomprehensible lyrics – what does that irritating line 'You are once, twice, three times a lady' actually mean? – is the most readable according to the tests.

There is a huge difference in scores between Flesch–Kincaid and SMOG. This may arise because SMOG wants a 100% comprehension level, Flesch–Kincaid only 75–85%. Based on the tests, the DfES declared that 'Angels' and 'Dancing Queen' required 'Level 2 skills', which it said are 'equivalent to the skills required to pass a GCSE at grades A* to C'.

People with level 1 skills are reading at the level of an average 11-year-old, said the DfES's press release, and: 'Level 1 skills would be required to pass a GCSE at grades D to G.' According to the DfES, 'people below this level may struggle to read these songs as well as check a pay slip for deductions and read bus or train timetables accurately'. The DfES says 5.2 million are below this level. (See appendix B for details of the levels and what they are said to mean.)

But the results rather depend on which readability test is used. Choose Flesch–Kincaid and the DfES won't get its press coverage because 'Angels' and 'Dancing Queen' score 9.1 and 7.9 years respectively, which are apparently within the range of level 1 readers. And a factor that the DfES ignores is that most people confident enough to do karaoke have memorized their favourite lyrics long before they enter the pub, so what's on the screen isn't that important.

Yet the DfES's research led the UK government's skills minister, Phil Hope, to opine: '...it's only on reading the lyrics properly that we realize that some of our favourite numbers are complicated. There are many pitfalls in public singing, but once you've got the mike in your hand you don't want your reading skills to let you down'.

In 2007, even more cash was spent on readability-test-based research which discovered that millions of people couldn't read and understand celebrity-chef recipes containing unusual ingredients like couscous and terms like 'blackened glaze' and 'aromatic shoulder of lamb' (Department for Innovation, Universities and Skills report, September 2007). The press salivated over the story, printing the usual gastro-porn pictures of Nigella Lawson's heaving décolletage as she sucked chocolate mousse off her fingers. The research, which proved only that unusual words – especially those of foreign origin – are, er, unusual was derided as a 'pointless official exercise in spending money' by Alice Miles, the *Times* columnist, in her piece 'Gobble-gobble. It's another Whitehall turkey' on 26 December 2007.

Is it good that government agencies are spending tax money on dodgy research whose sole purpose is to attract media attention? If the public really needs to know how many people can't read particular song lyrics or recipes, why not do a proper piece of research? It could use real people and compare their performance on lyrics and recipes with what they achieve when reading texts of known difficulty levels.

2 What readability tests won't tell you

Readability tests can be useful as propaganda when trying to convince people that their documents are too hard for the intended readership, and they also give a very rough idea of the level of difficulty. But they need to be used honestly, with caveats about what they don't do:

- They don't tell you whether the content is any good – whether the facts and

figures are right (although, to be fair, most plain-language editors won't do that either).

- They don't tell you whether the purpose of the document has been clearly stated.
- They don't tell you about the level of abstraction in the text – how nebulous or theoretical it may be.
- They don't tell you whether the text is logically arranged and whether it puts the big news early (if that's important).
- They don't tell you whether the headings and subheadings are good signposts pointing in clearly different directions and whether they form a sensible hierarchy – in fact, most readability tests require you to omit headings from the calculations.
- They don't tell you whether connector words are present to show the logical flow of the argument.
- They don't tell you whether there are personal-reference words like 'we' and 'you', which can often make the text more accessible.
- They don't tell you whether the grammar and punctuation are any good.
- They don't tell you whether the presentation is appealing or appalling.
- They don't tell you about readers' likely motivation levels.

In short, they don't tell you about a lot of things that are important. They are fixated on sentence length and syllable counting, which they make proxies for everything else that matters. (And SMOG doesn't really bother with sentence length – it's based mainly on the number of polysyllabic words.) Long sentences and unusual words can be serious problems for readers, we all know that. But so many other things matter, too, which readability tests ignore. The late Rudolf Flesch, a plain-English guru and originator of the readability tests that bear his name, is quoted approvingly by many who advocate the tests, and he said:

'Some readers, I am afraid, will expect a magic formula for good writing and will be disappointed with my simple yardstick. Others, with a passion for accuracy, will wallow in the little rules and computations but lose sight of the principles of plain English. What I hope for are readers who won't take the formula too seriously and won't expect from it more than a rough estimate.' (Quoted in *Readability Assessment of British Internet Information Resources on Diabetes Mellitus Targeting Laypersons*, School for Health, University of Bath, Sept 2004, by Maged Boulos)

But let's look further at what the tests don't tell us about. I received a standard letter recently from a pharmaceutical firm, which I'll call FBC Pharma. It starts:

'Dear Supplier

SAP IMPLEMENTATION

FBC Pharma Ltd is implementing SAP and has a "go live" date of 2 July 2007.

As we get accustomed to the new system, please be aware that response times may be slower than normal. As a result, we may not be able to meet payment terms but will endeavor to pay all invoices as close to the payment date as possible. Many thanks for your understanding in this matter.'

Most days I chuck things like this in the bin. That day I thought I'd ring up the signatory, whom I'll call Karen Burns.

- 'Hello, is that Karen Burns?'
- 'Yes.'
- 'I've got your letter about SAP implementation.'
- 'Oh yes?'
- 'What exactly is 'SAP'?'
- 'Er. Oh. I've no idea what it stands for – it's some German name.'
- 'But what does it stand for – what does it, you know, mean?'

- 'Well, it's a well-known accounting package. An enterprise-resourced package. It's a German thing.'

It was clear that Karen Burns, whose title was 'Finance Manager – UK Accounting and Reporting' had no idea what her 'enterprise-resourced package' was. The web offers this explanation:

'The original name for SAP was German: Systeme, Anwendungen, Produkte, German for "Systems Applications and Products". The original SAP idea was to provide customers with the ability to interact with a common corporate database for a comprehensive range of applications. Gradually, the applications have been assembled and today many corporations, including IBM and Microsoft, are using SAP products to run their own businesses. SAP applications, built around their latest R/3 system, provide the capability to manage financial, asset, and cost accounting... .'

My point is that if you put letters like FCB Pharma's through readability tests, which will of course require you to ignore the headings, then their short sentences and short words will make them seem easy for most readers.

I could multiply examples. I regularly see documents that would score well on readability tests but are useless for the intended readers, mainly because they make no sense if you try to work out what they mean instead of just looking at the words superficially. One example will do. It's taken from a booklet about how to fit a child's safety seat in cars, which scores reasonably well on the tests (ie, would be readable by most people). This bit is on the left-hand page of a 2-page spread:

'How to fit the Club Class Extra Rearward Facing
Have you checked that your baby weighs less than 13kg?
In this weight range you must only install the Club Class Extra rearward facing... .'

And then this is on the right-hand page:

'How to fit the Club Class Extra Forward Facing
Have you checked that your baby weighs more than 9kg?
In this weight range you must only install the Club Class Extra forward facing.'

These sentences, taken individually, are clear and will contribute to a good readability-test result for the whole booklet. But you'll notice that, taken together, they are nonsense because a child of 11kg appears to fit in both ranges. A readability test wouldn't tell you this was nonsense; quite the opposite.

The tests are blunt instruments. Look at this piece of text from the website of one of our local councils (name changed):

'An Ofsted inspection of its information and advice provision in Blankshire as long ago as 2005 remarked that they could not feel confident in the absence of support from our information and advice services that some vulnerable young clients would survive their transition to adulthood.'

Of course, that sentence would score poorly as part of a readability test. But it wouldn't tell you why people might have difficulty with it, apart from pointing out some polysyllabic words and the 45-word sentence. Nor would it suggest how to put it right. The apparent literal meaning is:

'Ofsted inspected our information and advice services in 2005. Ofsted thought it probable that if some vulnerable young people didn't get our advice and information, they would die before becoming adults.'

Now it probably wasn't supposed to mean that. When I put that translation to the local council that wrote it, they said they hadn't meant to say that. They couldn't actually remember what they had meant to say, but it wasn't that.

3 Are discussion groups more useful than readability tests?

In *Lucid Law* (Plain Language Commission, 1994 and 2000) I sought to discover whether a UK Act of Parliament rewritten and restructured in a plainer-English style would be easier for people to read and use. I hired a researcher to test both versions of the Timeshare Act (the real Act and my pretend one, the Clearer Timeshare Act)

using discussion groups of student lawyers and, separately, 'ordinary citizens'. She got each person to answer specific questions based on both Acts and compared their performance, as well as holding discussions to gauge opinions on the 2 versions – so it was a kind of usability test. The full research method, questionnaires etc are shown in the book (free download, <http://www.clearest.co.uk>, 'Books'). On one question, crucial to understanding the Act correctly, 94% of the student lawyers got the right answer when using my pretend Act but only 48% of them did so with the real Act.

The discussion-group results, based as they were on results with real people, were more persuasive than any readability test based on counting syllables would have been. The rewrite reduced the average sentence length by a third, reduced the overall length of the Act by a quarter, introduced far more headings, grouped the information more logically, and provided a full contents list. Apart from sentence length, these factors are ignored by most readability tests. The typography was also improved – again, the tests ignore this. The UK's Government-funded Basic Skills Agency sees typography as crucial:

'Difficulties with reading are often more to do with the look and layout of a text, than with the complexity of the text itself. In their efforts to produce attractive, eye-catching material, designers are sometimes tempted to sacrifice clear layout. Complicated layout and design can confuse those with reading difficulties.' ('Readability. How to produce clear written materials for a range of readers.' Basic Skills Agency booklet number A1880 – undated, probably 2005)

Desirable though discussion groups may be, however, they are not practical for most of the small-scale writing and editing jobs we tackle. They cost too much to convene and they take too long to set up and run properly. Most of our customers want a 2–3 day turn-round on their editing jobs and they don't want a huge bill. Like most plain-language practitioners, we therefore apply our judgment and experience to the task of writing or editing, without using either discussion groups or readability tests (though I accept that it is reasonable to take account of a readability-test result). It makes sense to convene discussion groups only when economy and time are not crucial factors. One-to-one interviews to check for comprehension are also likely to help. Interviewees' abilities can first be checked against passages of known difficulty to help make them representative of readers in a particular ability group.

4 The problem of irreducibility in plain-language work

EC directives

Bill DuBay, a proponent of what he has called 'the science' of readability testing, advocates putting European Commission laws through readability tests because he feels this would help to alert their authors to the level of difficulty in them.

Yet EC directives and regulations are inevitably full of polysyllabic words and many of these are irreducible. Examples include *bioavailability*, *directive*, *specimen*, *parliament*, *requirement*, *provision*, *effective*, *harmonized*, *regulation*, *standard*, *marketed*, *inspection*. Some are not particularly high-register or difficult but they ensure that laws on complex topics (that's most of them) will inevitably exceed 7th-grade, 8th-grade, 9th-grade or even 10th-grade levels as judged by readability testing. Bill DuBay suggests that drafters should be required to write law at the national average reading level (his PLAIN forum post, 12 September 2006) – which he takes to be 7th grade (UK reading age 12) but the inherently high-level vocabulary makes this impractical in most cases. Of course, the drafters could reduce sentence length much more than they do, but a test like SMOG is so fixated on syllable counting that this wouldn't help them to produce a good test score.

Edwin Tanner's point in the *Statute Law Review* (Vol 25, Number 3, 2004, p223), and mine in *Clarifying Eurolaw* (Plain Language Commission, 2001) to which Tanner refers, is that poor grouping of information, outdated legal-drafting conventions and abysmal sentence structure are the main factors that make EC laws and directives so hard to read – even for lawyers. Readability tests have nothing to say about these things. They just show that the stuff is very difficult. As I said in section 2, this is why they are so crude for assessing all documents for adults, not just EC laws.

The Clearer Timeshare Act has been criticized by readability testers for being '13th grade' (UK reading age 18), but this shows how facile the tests can be. The Act's average sentence length is 20–24 words, depending on how you treat vertically listed items. There are about 109 long words (ie, those of 3 syllables or more, ignoring repeat words) out of about 3,020 (excluding the citizen's summary). This seems a very low rate – only 3.6%. Here's a list of them:

authority, agreement, diligence, principal, prosecution, effectiveness, provisions, dictionary, enforcement, summary, purposes, prescribed, right-to-cancel, document, specified, considering, contravenes, entering, customer, cancellation, conviction, indictment, statutory, maximum, contravening, caravan, beforehand, circumstances, expiry, period, significant, accordance, unenforceable, enforceable, reasonable, reliance, another, permission, complying, proceedings, identify, information, default, liability, corporate, connivance, director, manager, secretary, capacity, liable, connection, management, committed, procedure, earlier, discovered, prosecutor, contrary, certificate, conclusive, interest, prominence, requirements, prescribed, legibility, unconditional, accompanied, regarded, properly, instalment, recalculated, repayment, intended, authorize, transaction, regulated, statutory, instrument, annulled, following, applying, financial, individual, intermittent, accommodation, anywhere, arrangements, insurance, employment, collective, verify, suspicion, relating, visible, legible, computer, incriminate, self-incrimination, paragraph, hindering, intentionally, material, obligations, necessary, typically, criminal, normally, protection.

How many of these long words could have been avoided without losing the legal effect and making the pretend Act ludicrous as a potential piece of law? For a topic as inherently complex as timeshare, the vocabulary was, I believe, close to its irreducible minimum. If someone is motivated to read the Act and is prepared to study it a bit, perhaps with the help of a dictionary, then I think they'd be able to get a reasonable understanding of most of it. That's what our testing with discussion groups showed.

Incidentally, the Clearer Timeshare Act uses 'agreement' repeatedly – because it's all about agreements – which readability tests would count as long or 'difficult' every time it appears, giving an artificially high score. Had the pretend law used 'contract' or 'thing that you do to make a deal', maybe the score would have been better. Someone recently suggested, perhaps tongue in cheek, that because 'electricity' was a hard word according to the testers, we should use 'power that you get from a plug socket'. But if that's anyone's idea of plain language, they can keep it.

So are readability testers really saying that a text with an average sentence length of some 20 words and 3.6% of long words puts it beyond the powers of 70–80% of Americans and Britons? If so, then there's been an even more catastrophic failure in our education systems than most of us dare to think.

Daily Mail's readability

And I do wonder how the popular press keeps selling so well. In the UK, the best-selling mid-market tabloid paper is the *Daily Mail*, with a daily sale of 2.4 million in December 2005 (compare the UK's top-seller *The Sun* with 3.1 million, and *USA Today* with 2.2 million). You can multiply these figures by, say, 4 to get the real daily readership. One of the *Mail's* best-paid columnists is Richard Littlejohn. His main piece on 12 September 2006 (appendix C) had some 860 words, an average sentence length of 20, plenty of passives, and about 100 long words (11%).

Assuming the long words are spread randomly, this gives a FOG score of about US grade 12.3 (UK 17.3), which is close to undergraduate level. Here's a list of them:

exemplary, exquisite, anniversary, million, Islamic, terrorists, preposterous, self-aggrandising, secretary-general, Islamophobia, widespread, dignifying, reaction, especially, regularly, authentic, victimhood, shtick, broadcasters, seriously, government, ministers, senior, officers, opportunity, agenda, utterly, terrorism, understand, exactly, responsible, excitable, community, criticize, Islamophobe, self-appointed, ridiculous, lookalike, officer, accusing, remotely, construed, religion, Christianity, political, certainly, bigotry, radical, enlightened, homosexuals, fantasies, smear-the-messenger, potential, anything, examine, guaranteed, suspicion, incendiary, occasion, reflection, remembrance, atrocity, perfunctory, condemnation, terrorism, caveat, assertion, smithereens, alienation, radicalized, precisely, fanatical, elements, maniac, undoubtedly, discrimination, unemployment, exclusive, alienated, particularly, ethnic, minority, justify, violent, politicians, absolutely, devices, integration, co-existence, purport, allegiance, Parliament, inflammatory, interventions.

So here is a popular columnist whose vocabulary is more rich and complex than appeared in the Clearer Timeshare Act – and far more so than appears in most government information for the public – and he uses 3 times more difficult words per 1,000. It's a miracle he sells stacks of newspapers, but he does. What does this tell us about readability tests? Perhaps that in the real world of writing for adults they don't really test what they say they do.

Using the best words for the job

Generally I think it's better to equip people with the proper words that professionals use about their specialisms – when these are essential to the message – than to pretend they use primary-school vocabulary. That's why I use 'trafficator' or 'indicator' not 'winker' (for goodness' sake). That's why, in leaflets I've edited recently, I've used (with explanations, glossaries etc) diminishing musharaka, ijara (types of Islamic mortgage), endoscopy, diabetic retinopathy, incapacity benefit, anaesthetist and vasovagal syncope. It plays havoc with the Flesch result but it is surely the right thing to do. What's more, I find it works. Some of the roughest toughest housing estates in Britain give favourable evaluations of the clarity of leaflets edited by our staff. Some stuff is just plain irreducible. I think we fool ourselves if we don't recognize that.

5 The notion of 'average reading age' (UK) and 'average grade levels' (US)

Differing views on the UK average

In *Lucid Law*, I said I wanted the law to be accessible to people of average intelligence and average reading ability. That was a bit woolly but I think it was realistic. A new law is usually about a complex thing and has to carve out a space in an existing body of law. This means it will rarely be feasible to write it all in simple, declarative, active-voice sentences of 15-20 words, despite the best efforts of good writers. So to most 7th-grade (UK reading age 12) readers it will always be a closed book, though authors should certainly do everything they reasonably can to keep the readability level down.

In his PLAIN forum post on 12 September 2006, Bill DuBay said: 'In the U.S., a full half of the people read below the 7th grade level. That is not surprising when we consider a quarter of the population has not graduated from high school. That group reads at the 3rd-grade level. Nine percent (that's 27 million) do not read at all. The stats for the U.K. are similar.'

The figures for the US may be correct (though I would question the inference some may draw that we must all therefore write at the 3rd-grade level (UK reading age 8)),

but my research suggests it is wrong to say that 'the stats for the UK are similar'.

The average UK reading age is hard to pin down. The UK's Plain English Campaign Limited (PEC), a well-known company in the field, has stated repeatedly (eg, *Mail on Sunday*, 3 March 2002) that the average reading age in the UK is 9.5 years (about US 4th grade), though when asked in writing for evidence PEC has refused to provide any. This level, of course, is below functional literacy, so PEC is asserting that the average UK adult is illiterate – an extraordinary state of affairs, were it true, given the amount of taxation pumped into our schools. But of course it isn't true.

Maged Boulos of the University of Bath, in his study of websites about diabetes mellitus (cited above) makes a similar claim to that of PEC, saying the 'estimated reading age of the UK population in general is 9 years', which is US 4th grade. He does not say who estimates it. Believing in this '9 years' average, he concludes from Flesch Reading Ease and Flesch–Kincaid testing alone that nearly all the websites in his sample exceed UK adults' average reading ability. This he asserts as fact.

Differing views on rates of illiteracy

In an email to me (27 January 2006), Maged Boulos quotes one of his main sources for the average adult reading age as an article on a BBC Northamptonshire website. Its author is a student whose web page includes the expressions 'boistorious', 'references', 'friends name' and 'friends address'. The article was a competition entry, and it didn't win. It said that the world has '810,879 billion people who are illiterate'. (The planet suddenly feels very crowded, with a terrible surfeit of libraries.) It also asserted that 20% of UK adults are illiterate: '...eight million people are so bad at reading and writing that they cannot cope with the demands of modern life'. This view is held by many, like the author of the popular BBC drama *The Amazing Mrs Pritchard*, who had one of her characters declare on 10 October 2006 that 7 million UK adults were 'functionally illiterate'.

Can this be right? Brace yourself for some complex figgerin', because the official statistics on UK reading ability are complex and poorly expressed. The National Literacy Trust (quoting a 2003 UK Department for Education (DfEE/DfES) study) says 5.2 million UK adults (16%) have reading skills at 'entry level 3 or below', though most of these (11%) are at entry level 3 itself. This (entry level 3) corresponds to 'the level expected of an 11-year-old' – which I take to be on the brink of functional literacy. The remaining 84%, however, are at higher levels, namely level 1 (40%) and level 2 (44%). See appendix B for more on the levels.

Actually there's total confusion between the National Literacy Trust's view of an average 11-year-old's ability (entry level 3, it says) and the DfES's view (level 1, says the press release mentioned in section 1 of this paper).

The Moser report, on which much current UK public policy on adult reading abilities is based, is full of guesswork and confusing calculations about the number of people with reading difficulties. It asserts that:

'Some 7 million adults in England – one in five adults – if given the alphabetical index to the Yellow Pages [a phone book listing businesses by type], cannot locate the page reference for plumbers. That is an example of functional illiteracy. It means that one in five adults has less literacy than is expected of an 11-year-old child. These figures – based on official surveys – are inevitably estimates, and may be a little on the high side: but the order of magnitude is certainly right.' (The Moser report, properly known as *A Fresh Start – improving literacy and numeracy* (DfEE 1999, ref: CMBS 1))

Having given the impression that functional literacy is what an average 11-year-old child would have, Moser then seems to suggest that it categorizes both the entry-level 3 group of adults (supposedly equivalent to average 11-year-olds) and those

below that level as functionally illiterate. For a report about good reading and writing, it is not at all clear.

DfES figures (<http://www.literacytrust.org.uk/socialinclusion/adults/survey.html>) show that only 5% of adults are at entry level 1 or 2 (that's only around 1.7 million). So I regard the DfES data as showing that in the UK 95% of adults are functionally literate, with the distribution strongly skewed towards higher reading abilities than mere functional literacy.

On top of its claim that one in 5 grown-ups in the UK can't find 'Plumbers' in the *Yellow Pages* index, the Moser report also says that one in 16 adults, if shown a poster advertising a concert with only 19 words on it of which one is 'Birmingham', cannot tell where the concert is being held! Apparently this is based on studies with real people – which should be the best method. I have 3 responses to this. The first is that I don't believe it. The second is that if it's true, then many schools have been failing to teach huge numbers of people to read the most basic words for some 50 years. The third is that I wonder if there is any point trying to write public information for people who can't find the word Birmingham in a 19-word poster and who may, in some cases, have decided that they'd prefer not to read at all – what may be called voluntary illiteracy. And I feel sure that however I write for them, they won't (a) pick it up because they just don't want to read, or (b) get beyond the first few sentences if they see a more complicated word than Birmingham, like 'Leicestershire' or 'Cirencester'.

Is 'average reading age' a useful idea?

Not everyone thinks the notion of 'average reading age' is useful. The Basic Skills Agency tells me (email, 27 January 2006): 'We don't actually work in terms of average reading ages at the Agency, we use a readability formula called SMOG, which helps to match the reading level of written material to the reading with understanding level of the reader.' The Agency's booklet (cited above, page 6) recommends a SMOG rating of under 10 for public information, saying: 'A readability level under about 10 will be able to be understood by most people.'

According to the SMOG test given in that booklet, this level 10 equates to no more than 2–3 words of 3 or more syllables every 10 sentences, no matter how long those sentences are! We may as well be writing nursery language if this is so. Using the booklet's SMOG test, the first 10 sentences of Richard Littlejohn's article, cited above, score 18. Unreadable by most adults in the country, supposedly, but they keep on buying the *Daily Mail*. The first 10 sentences of the Clearer Timeshare Act score 13 if repeat words are ignored (which the tests don't seem to advocate). That seems to be roughly the 7th or 8th grade – UK reading age 12 or 13.

Who should be included in the average?

Should we include non-readers in the calculation of average adult reading ability? Is it sensible to include the one million eastern Europeans who have migrated to the UK since their countries acceded to the EU 3 years ago, in a calculation of the average ability of UK adults to read English? If you also remove from the calculation of 'average' the 2 million people said by the Foundation for People with Learning Disabilities to have 'learning disabilities' (www.learningdisabilities.org.uk), then you have a somewhat higher reading age or grade level to aim at for your public information. Perhaps, given the complexity of some of the topics, this higher target would make more sense than what many authors are haplessly trying to do now,

which is to write booklets about, say, divorce and separation law for people who can't read 'Birmingham' in a 19-word advert.

Maybe foreign non-readers of English would be better relying on translations in their own languages such as Polish or Urdu – which in the UK are often available – as well as learning English. British non-readers of English would perhaps be better using word-of-mouth communication including tape and video. (Tape is often available from public bodies, and other formats are also heavily advertised.) I don't know about the US, but maybe many non-readers of English in the US are literate in Spanish or some other language – so I assume public information is made available to them in a tongue they can read. This would seem to make more sense than including them in the figures for illiterate native English speakers and trying vainly to adapt English-language materials to their needs.

What might the average be?

I asked a mathematician to look at the figures on the (UK) National Literacy Trust website (appendix B). He deduces that the average reading age of the entire sample is 14.5 years; or, if entry levels 1 and 2 (the 5% of the sample who are non-readers) are disregarded, then the average is 15 years. So if we err towards pessimism, it seems that the average UK adult has a reading age of perhaps 12–14.

What does an average of 12–14 look like?

Bill DuBay has kindly sent me some samples of graded texts ('normed passages') for schoolchildren. I think they're very useful in showing what kind of level authors should hit when writing for a mass audience. Here's an example from a school textbook that is supposed to be grade 8 (UK 13 years):

'As we have seen, a neutron star would be small and dense. It should also be rotating rapidly. All stars rotate, but most of them do so leisurely. For example, our Sun takes nearly one month to rotate around its axis. A collapsing star speeds up as its size shrinks, just as an ice-skater during a pirouette speeds up when she pulls in her arms. This phenomenon is a direct consequence of a law of physics known as the conservation of angular momentum, which holds that the total amount of angular momentum in a system holds constant. An ordinary star rotating once a month would be spinning faster than once a second if compressed to the size of a neutron star. In addition to having rapid rotation, we expect a neutron star to have an intense magnetic field. It is probably safe to say that every star has a magnetic field of some strength.' (From *Discovering the Universe* (Faufmann 1990, p. 290). Quoted in W H DuBay *The Principles of Readability*, quoting 'Qualitative Assessment of Text Difficulty, A Practical Guide for Teachers and Writers' by Jeanne Chall and others (1996). Two of Bill DuBay's other books on readability tests are *Unlocking Language* (2007), Impact Information, Costa Mesa, USA, ISBN-4196-6176-0 and *Smart Language* (2007), same publisher, ISBN 1-4196-5439-X.)

In fact the tests give varying results: Flesch–Kincaid US grade 9, SMOG 11.7 and FOG 12.9. Let's see what the grade scores would be if we put the sentences in reverse order:

'It is probably safe to say that every star has a magnetic field of some strength. In addition to having rapid rotation, we expect a neutron star to have an intense magnetic field. An ordinary star rotating once a month would be spinning faster than once a second if compressed to the size of a neutron star. This phenomenon is a direct consequence of a law of physics known as the conservation of angular momentum, which holds that the total amount of angular momentum in a system holds constant. A collapsing star speeds up as its size shrinks, just as an ice-skater during a pirouette speeds up when she pulls in her arms. For example, our Sun takes nearly one month to rotate around its axis. All stars rotate, but most of them do so leisurely. It should also be rotating rapidly. As we have seen, a neutron star would be small and dense.'

Of course, the scores would be virtually the same, even though the paragraph is now poorly structured and nonsensical. This is because the tests do not recognize poor structure – most of what they do is counting.

The following normed passage is also said to be grade 8 (UK, 13), although it seems far simpler than the neutron star text:

'While all this is happening, the embryo is getting longer. Now the tail bud begins to develop, and the

embryo develops suckers beneath the place where the mouth will be. Although it is only about twenty-eight hours old and barely recognizable as a tadpole, the embryo now hatches. Toad embryos emerge very early, while frog and salamander embryos are further along when they come out of their jelly prisons. The small embryos hang by the suckers to the jelly. It will be five more hours before they can move their muscles at all.

When they are a little more than an day and a half old, their hearts begin to beat. Soon the blood begins to circulate through the gills, the developing eyes can be seen, the mouth opens, and the suckers begin to disappear. At two and a half days of age, when the blood starts circulating in the tail, they really look like tadpoles.'

Flesch–Kincaid scores it 7.5, SMOG 8.5, and FOG 9.3. These scores are indeed much lower than those given for the neutron star text. It's puzzling that 2 grade-8 texts should have such differing scores.

The main problem with readability tests

Normed passages tend to have one thing in common: they are basically well written. They tend to be clear, logical and well structured. When a readability test like SMOG or Flesch–Kincaid is applied to them, they will get a certain score. And when the same test is applied to an ill-structured or otherwise incoherent piece that has similar properties (eg, vocab, sentence length), it will get a similar score. This is the basic problem with tests like SMOG. Those who apply them uncritically tend to assume that any 10-sentence passage with, say, 12 polysyllabic words is as good and clear as any other with 12 polysyllabic words. But its grammar and punctuation may be poor and its message muddled, ambiguous or misleading. Such findings are only likely to emerge after usability testing (not readability testing) or editorial scrutiny or both.

6 Are plain-language principles relevant to works of literature?

In my view, plain-language practitioners should confine their comments to stuff that's meant to be essential public information (like law texts, government information, school textbooks, and customer information) rather than broadening them to newspapers or creative stuff like novels, poetry, song lyrics, and the broadcast spoken word. If the purpose of the document or website is clearly meant to be public information, then it's fair game. Otherwise I think we should leave it alone. There is a risk that in calling for the same kind of clarity everywhere – eg in celebrity-chef recipes – we'll deny people exposure to words they may like to know and words they may find express their thoughts and feelings well. Words are not just for clarity – they also resonate, they carry baggage, and they are for pleasure, play and sport. There's a dictionary full of words and I don't think we should be suppressing any of them on the grounds that they'll frighten weak readers. The only way people learn new words is by encountering them. All of us need to know more words, not fewer. How people can be clear when communicating *essential information* is our province, and we should not stray far outside it.

7 Coming soon: a resource for authors and editors

Some readability tests are based on lists of words said to be understood by readers at particular US grade levels, notably the *Living Word Vocabulary* (LWV) by Dale and O'Rourke (1979). I've compiled a vocabulary guide for plain-language authors and editors that includes words from the LWV alongside their grade level as calculated by Dale and O'Rourke's face-to-face testing, followed by a UK reading level, followed by a frequency based on the British National Corpus of 100 million words. There are also plainer alternatives for certain words, and some commentary. This 'Plain English Lexicon' will be on our website early in 2008.

Appendix A: the song lyrics we tested

The lyrics were tested by my colleague Sarah Carr in September 2007 using Micro Power & Light software. They were put in to a text layout for the purpose of the test, with line returns removed from each verse. The only punctuation added was full stops. Semicolons, dashes etc would not have affected the scores so they were not added. Run-on sentences and missing apostrophes are immaterial to the scoring so these and other grammatical errors were not corrected. Where a line ended in a comma, it was not changed to a full stop.

'Angels'

I sit and wait. Does an angel contemplate my fate. And do they know the places where we go when we're grey and old 'cos I've been told that salvation lets their wings unfold. So when I'm lying in my bed thoughts running through my head and I feel that love is dead I'm loving angels instead.

And through it all she offers me protection a lot of love and affection whether I'm right or wrong. And down the waterfall wherever it may take me I know that life wont break me. When I come to call she wont forsake me. I'm loving angels instead.

When I'm feeling weak and my pain walks down a one way street I look above and I know ill always be blessed with love. And as the feeling grows she breathes flesh to my bones. And when love is dead I'm loving angels instead.

And through it all she offers me protection a lot of love and affection whether I'm right or wrong. And down the waterfall wherever it may take me I know that life wont break me. When I come to call she wont forsake me. I'm loving angels instead.

'Dancing Queen'

You can dance, you can jive, having the time of your life. See that girl, watch that scene, dig in the dancing queen.

Friday night and the lights are low. Looking out for the place to go where they play the right music, getting in the swing you come in to look for a king. Anybody could be that guy. Night is young and the musics high. With a bit of rock music, everything is fine. Youre in the mood for a dance. And when you get the chance...you are the dancing queen, young and sweet, only seventeen. Dancing queen, feel the beat from the tambourine. You can dance, you can jive, having the time of your life. See that girl, watch that scene, dig in the dancing queen.

Youre a teaser, you turn em on. Leave them burning and then youre gone. Looking out for another, anyone will do youre in the mood for a dance. And when you get the chance...you are the dancing queen, young and sweet, only seventeen. Dancing queen, feel the beat from the tambourine. You can dance, you can jive, having the time of your life. See that girl, watch that scene, dig in the dancing queen.

'Three Times A Lady'

Thanks for the times that you've given me, the memories are all in my mind. And now that we've come to the end of our rainbow there's something I must say out loud.

You're once, twice, three times a lady, and I love you. Yes, you're once, twice, three times a lady, and I love you. I love you.

When we are together the moments I cherish, with every beat of my heart. To touch you, to hold you, to feel you, to need you. There's nothing to keep us apart.

You're once, twice, three times a lady, and I love you. I love you.

'White Room'

In the white room with black curtains near the station. Blackroof country, no gold pavements, tired starlings. Silver horses ran down moonbeams in your dark eyes. Dawnlight smiles on you leaving, my contentment.

I'll wait in this place where the sun never shines; Wait in this place where the shadows run from themselves.

You said no strings could secure you at the station. Platform ticket, restless diesels, goodbye windows. I walked into such a sad time at the station. As I walked out, felt my own need just beginning.

I'll wait in the queue when the trains come back; Lie with you where the shadows run from themselves.

At the party she was kindness in the hard crowd. Consolation for the old wound now forgotten. Yellow tigers crouched in jungles in her dark eyes. She's just dressing, goodbye windows, tired starlings.

I'll sleep in this place with the lonely crowd; Lie in the dark where the shadows run from themselves.

Appendix B: National Literacy Trust figures on adult literacy

Information from the Trust's website in October 2007.

Literacy results

Level	% of 16-65 yr olds	Number of 16-65 yr olds
Entry level 1 or below	3%	1.1 m
Entry level 2	2%	0.6 m
Entry level 3	11%	3.5 m
(All Entry level 3 or below)	(16%)	(5.2 m)
Level 1	40%	12.6 m
Level 2 or above	44%	14.1 m

Base: all respondents with literacy score (7,874)
Source for population figures: Census 2001

More than a third of adults living in social class five households were classified as having lower literacy skills, whereas adults living in social class one households were roughly four times more likely than those in social class five households to reach Level 2 or above in the literacy test.

Language was a barrier to those whose first language was not English (seven per cent of the total) - only one in four achieved Level 2 or above in the literacy assessment. However, those whose first language was not English but who claimed to have 'very good' spoken English performed to a similar standard as those with English as a first language. Among those speaking English as their first language, there were only minor differences in skills levels between the various ethnic groups. The exception was the low performance of the English speaking Black Caribbean population.

Good literacy and numeracy skills tended to be associated with good wages, although the connection was stronger for numeracy.

Very few adults regarded their reading, writing or maths skills as below average, even among those with the lowest level of ability. Only a tiny proportion (two per cent) felt their weak skills had hindered their job prospects or led to mistakes at work.

Literacy and numeracy skills in everyday life

The majority of respondents at each level of literacy claimed to read every day, with the exception of Entry 1 or lower level literacy. One in four of these respondents said they never read, but even among this group, four in ten read every day. The frequency of writing was more closely associated with literacy level. Only one in five of those with Entry 1 or lower level literacy, and only one in three of those with Entry level 2 literacy wrote every day.

Nearly all parents of children aged 5-16 said they helped their children with reading (95 per cent), writing (89 per cent) or maths (87 per cent). Those with lower levels of literacy and/or numeracy were less likely to help, and were less confident when they did, but still the majority tried to help (63 per cent of those with Entry level 2 or lower level literacy helped with their children's reading).

Adult literacy levels - what do they mean?

Very few adults can be described as 'illiterate' - most will have some level of reading or writing skills, and strengths or weaknesses in particular areas. The Skills for Life adult basic skills strategy, launched by the Government in 2001, developed national standards for literacy, numeracy and ICT. Each framework outlines what an adult should be able to achieve at entry level (divided into three sub-levels), level 1 and level 2 or above. The frameworks recognise that an adult may be classified at an overall level of skill but have higher or lower levels of ability in different aspects of that skill - what has been described as an uneven or 'spiky' profile. The framework for literacy is outlined below. It should be noted that skills levels and tests for adult literacy refer primarily to reading skills, rather than writing.

Level	Literacy (reading) An adult classified at this level...	Equivalent to...
Entry level 1	- Understands short texts with repeated language patterns on familiar topics - Can obtain information from common signs and symbols	National curriculum level 1
Entry level 2	- Understands short straightforward texts on familiar topics - Can obtain information from short documents, familiar sources and signs and symbols	Level expected of a seven-year-old (national curriculum level 2)
Entry level 3	- Understands short straightforward texts on familiar topics accurately and independently - Can obtain information from everyday sources	Level expected of an 11-year-old (national curriculum levels 3-4)
Level 1	- Understands short straightforward texts of varying length on a variety of topics accurately and independently - Can obtain information from different sources	GCSE grades D-G (national curriculum level 5)
Level 2	- Understands a range of texts of varying complexity accurately and independently - Can obtain information of varying length and detail from different sources	GCSE grades A-C (national curriculum levels 6-8)

These levels descriptors are taken from *The Skills for Life Survey: a national needs and impact survey of literacy, numeracy and ICT skills*, published by the Department for Education and Skills in 2003.

Appendix C: Littlejohn column, Daily Mail, 12 Sept 2006

This box is blank because copyright permission is awaited.

