

What makes a good document?

The criteria we use

Rob Waller *April 2011*

During our first two years, one of our services to member organisations was benchmarking – this means evaluating their documents using a common set of criteria, so they can see how theirs compare to others.

Our benchmarking process is described in *Technical paper 5: Benchmarking everyday documents*. This paper discusses the reasoning and research evidence behind our criteria in more detail.



This paper can usefully be read alongside three others in this series:

- 5 Benchmarking everyday documents
Summarises our first nine benchmarking studies.
- 7 What do people notice about their documents?
We asked members of the public to keep document diaries recording their reactions.
- 8 Criteria for clear documents: a survey
The criteria used by various appraisal schemes around the world.

They can be downloaded from www.simplificationcentre.org.uk/Resources/ or from www.reading.ac.uk/cidr

Our benchmarking process

Benchmarking means comparing the quality of documents from different organisations. Our benchmarking reports are structured expert reviews that pinpoint strengths that can be celebrated and weaknesses that need to be addressed.

During our first two years of operation we were funded through a membership scheme, and organisations who joined the Simplification Centre could take part in our benchmarking exercise. Benchmarking is a process commonly used in business to establish a performance standard for organisations to aspire to. Rather than compare with an absolute standard of perfection, which may be hard to define and impossible to reach, in benchmarking your performance is compared with other organisations. This helps you understand how much better you might expect to perform, and, if information is available about how other organisations achieve their standard, it might also point the way to more effective business processes.

There are various different ways to carry out benchmarking, including user testing, collection of actual performance data, and expert review. For practical and cost reasons, we use expert review, with a simple scorecard system. In a separate document (*Technical paper 5: Benchmarking everyday documents*) we report on what we found in the first set of nine benchmarking exercises we have carried out using the criteria we describe here.

We designed our benchmarking scorecard to reflect best practice in document design. Our idea of best practice is supported by various different kinds of knowledge: our own experience as document designers, our own research programme collecting customer feedback on documents, and academic research published over many years, from a variety of disciplines, such as psychology, linguistics and educational theory. This review summarises some of the reasoning, theories and research findings that underlie our criteria.

Our criteria cover a very wide range of factors, and so potentially call into play a vast literature from a several different disciplines. We are aware of a certain foolhardiness in trying to cover such a wide field, and stress that this document is a work in progress, a reflection on practice that is pragmatically, not ideologically or theoretically motivated.

Some key concepts

Certain themes underpin many of the criteria and guidelines we use in the Simplification Centre.

One is the notion of **cognitive capacity** (related terms you may hear are ‘cognitive load’ or ‘performance load’). Psychologists theorise that there are limits on the amount of information that can be processed at one time. Although it has been superseded now, theoretically speaking, a famous paper by the psychologist George Miller (1956) identified the ‘magic number seven, plus or minus two’ as the number of categories that can be easily handled in working memory. Many design and writing guidelines are there to ensure that working memory does not become overloaded. For example, a long sentence with complex clauses requires more cognitive capacity to process than a short simple sentence. Research has shown that people read difficult text more slowly, because they have to make more effort to decode words, recall or infer difficult meanings, and maintain a large number of new concepts in working memory (Petros, Bentz, Hammes & Zehr 1990).

Another key theme is **strategic reading** (Paris, Wasik & Turner 1991). Readers are not passive sponges, soaking up information as it is fed to them line by line. The most effective readers are aware of their objectives, monitor the relevance of each part of a document to those objectives, and select the most relevant parts to attend to. Readers need surface level cues to help them do this effectively. So we look for documents to signal their purpose, intended audience, context, and structure.

A third theme, that underlies a number of the criteria, is **affordance** (Gibson 1977). This term refers to the quality of a design that allows or even encourages certain kinds of user behaviour. A classic example is of a door handle (it not only allows you to open the door, but its design can tell you which way the door is hinged, and whether to push or pull – in many cases it positively invites you to act in the intended way). The equivalent qualities in documents might be large print summaries that allow and encourage (ie, afford) previewing of key content, checklists that afford the correct returning of key documents such as forms or payments, and contact information that affords the use of the right channels for customer queries.

The concept of affordance works at different levels. At the level of perception, we actively seek meaning in the world around us – in a document we tend to see significance in the way things are aligned, and in their relative prominence. At the level of interpretation,

we look for codes, conventions, implications and other kinds of significance. And at the level of navigation, we want to see larger structures and organising principles to help us decide whether and how to read.

Another important theme is the concept of **schemata**, and theoretical variants such as mental models, or scripts (Kintsch 1974, McNamara et al 1984; McVee et al 2005). Cognitive psychologists have established that our understanding of verbal information draws heavily on pre-existing knowledge or frameworks, sometimes referred to as schemata. These are conceptual structures, sets of expectations, or mental scripts that we can use to make inferences that may not be explicit in the text itself. Scripts are sets of knowledge about what we expect certain situations to be like and what might normally happen in them, based on experiences we have gathered over time. A much used example is the restaurant script: if we say ‘Nick went into the restaurant. He ate a steak. He left’, people will bring a wide set of elaborations to the story from their shared contextual knowledge – ie, that he sat down, read the menu, paid the bill, etc.

Most of us have good schemata for everyday events like meals, that we can rely on until we go to a foreign country where things are done differently. But few of us have an equally good schema¹ for choosing a pension, and people entering adulthood may have only a sketchy idea of what is involved in renting a house, paying tax, buying insurance.

.....

1 Following the original Greek, the singular is ‘schema’, the plural ‘schemata’.

The sixteen criteria

We benchmark documents against sixteen criteria, which fall into four broad categories.

Language criteria	How easy it is for people to understand the words	
<i>Directness</i>	Using direct language to make clear who's doing what.	6
<i>Plain words</i>	Extent to which the vocabulary is easily understood.	8
<i>Grammar and punctuation</i>	Conformity with the practice of good standard English.	10
<i>Readability</i>	Ease with which the reader can follow the argument of the text.	12
Design criteria	The visual impact of the document and the way its design influences usability.	
<i>Legibility</i>	Use of legible fonts and text layout.	15
<i>Graphic elements</i>	Use of tables, bullet lists, graphs, charts, diagrams, etc.	16
<i>Structure</i>	Quality of the document's organisation in relation to its function.	17
<i>Impression</i>	Attractiveness and approachability of the document's overall appearance.	18
Relationship criteria	How far the document establishes a relationship with its users	
<i>Who from</i>	Is it clear who is communicating?	20
<i>Contact</i>	Whether there are clear contact points and means of contact.	21
<i>Audience fit</i>	Appropriateness to the knowledge and skills of the users.	22
<i>Tone</i>	Matching the style and language to the context.	23
Content criteria	How the content and the way it is organised deliver the document's purpose	
<i>Relevance</i>	How relevant the content is to the recipient.	25
<i>Subject</i>	Whether it is clear what the communication is about.	27
<i>Action</i>	Clarity about what action is required of the user.	28
<i>Alignment</i>	Compliance with the organisation's intended aims and values.	29

1

Design criteria assess the visual impact of the document and the way its design influences usability.

“They should get to the point, cut to the chase. What action is there to take, and what are the implications?”

These marginal quotes are from document diaries kept by members of our research panel, or from follow-up interviews.

1.1 Directness: using direct language to make clear who’s doing what

This criterion groups together several aspects of clear writing that have to do with getting to the point, and making it clear who the participants of actions are.

Reasoning and evidence

When we read or hear a sentence, we tend to expect to be able to find a ‘who did what to whom’ structure or some variant of this.

Passives are a well known cause of indirectness in language. A passive sentence is one in which the doer (the person doing the action, as in ‘Tom kicked the ball’) is hidden (as in ‘The ball was kicked’). The opposite of a passive sentence is an active sentence.

When we read written language, we derive only part of our interpretation from the words in the text – the rest is supplied by us from our general knowledge, or from inference (Kintsch & van Dijk 1978).

For example, the passive sentence ‘A refund will be sent’ implies that someone will send it and someone will receive it, but neither party is identified. ‘You will be sent a refund’ identifies the recipient, while ‘we will send you a refund’ is even more informative because it identifies the sender as well. Because English is very strongly a Subject Verb Object (SVO) language, its speakers have the expectation that the subject (the first thing they encounter) is typically the doer of the action. Passive constructions are harder to process because the reader has to work out which of two possibilities (the ‘you’ and the missing doer, in this case ‘we’) is the correct solution (Slobin 1966). This is a sub-conscious process that creates an additional load on short-term memory whenever there are two possible doers of an action. Passives are also less frequently used than actives (and so less expected) and children acquire them much later – another popular measure of complexity (Horgan 1976).

In everyday topics, this may not matter because the reader can reliably be expected to know how the world works. And because we assume the writer is cooperating with us in a sincere effort to communicate something that is relevant to us (Grice, 1975, developed what he calls the Cooperative Principle), we are usually very good at inventing ‘possible worlds’ in which even incomplete sentences make sense.

In specialist areas such as personal finance or tax, however, we cannot assume that the customer has very much prior knowledge (see the discussion of financial capability, under the *Audience fit* criterion). There is a range of participants in every process (for example, a pension plan may involve the customer, the customers’ dependants or beneficiaries, the financial adviser, the insurance company, the tax office and the regulator), and they operate specialised processes that ordinary people cannot be expected to understand. They want full explanations, and active sentences help to prevent writers from leaving important things unsaid.

Nominalisations are another source of indirectness. This term describes the use of a noun to refer to an action (for example, ‘the determination of your bonus’, or ‘correspondence should be addressed to...’), leaves the reader to infer who the participants might be, in a similar way to passives. To fully understand the meaning, readers must at least infer the agent of the action (‘We will determine your bonus’, ‘If *you* wish to contact us...’). This again adds to processing load. In some cases, the reader also has to work out whether the noun refers to an event at all: for example, ‘publication’ could be the action of an organisation publishing reports, or an actual document. The use of such nominalisations is widely criticised in guides on clear writing.

Exceptions

As with every rule, there are exceptions. Some researchers have demonstrated that there are cases where passive constructions are preferable to active ones (Hupet & LeBouedec 1975): for example, where you want to draw special attention to the done-to person. For example: ‘Your financial adviser will be sent the refund’. In these cases, it can be even clearer to say: ‘It is your financial adviser who will be sent the refund’. Ferreira (2003) found that constructions like this (known as cleft sentences) were as quickly and accurately processed as active sentences, and more quickly processed than passive ones.

What we look for

We check for the use of the active voice where possible; the first rather than third person; compact expressions ('Send back') rather than indirectness ('We would appreciate it if you would send back'); expanded noun phrases ('How much your pension might be worth in the future') rather than nominalisations ('your projected fund')

"There's jargon. 'Returns of financial stocks were more muted.' I can guess what it means, but how a return can be muted I'm not sure. Presumably returns means money back, or money being paid out, so I think it could be linked to money, I'm not sure. And here 'market sentiment had become muted'. They obviously like the term. Perhaps people weren't yelling so loudly as before! So I do think that perhaps for the less financial of us it's difficult"

1.2 Plain words: the extent to which the vocabulary is easily understood

The use of simple, common words is at the heart of all plain English advice. We may not always be able to entirely avoid specialist language, but our instinct should always be to use the shortest, commonest and most expected word.

Reasoning and evidence

In the research literature there are various reasons why words are considered difficult. These include:

Familiarity The English language has many thousands of words (the OED contains as many as 615,000) but the average person only uses a small fraction of them (estimates vary widely, between 15,000 and 50,000), and this has been attributed to the amount of reading they do (Nagy & Herman 1987).

Common words are easier to understand than uncommon ones (Wheeler 1970). Research reviewed by Felker et al (1981) showed that uncommon words affect the time it takes to read sentences, they make it harder to remember them, and they also affect meaningful tasks such as judging sentences to be true or false. It has been suggested as an explanation that familiar words or concepts can be more easily referenced from long-term memory, and working memory is freed up for processing new information.

Historically, English developed from two main sources – Scandinavian or Anglo-Saxon languages spoken by invaders from Northern Europe, and Norman French, a Latin-based language, and Latin itself. Because of our history, Latin words are associated with government, bureaucracy and the law, and are less common in everyday speech. So when people talk about plain English, they often mean using the Anglo-Saxon word, not its Latin equivalent – 'get' instead of 'receive', for example, or 'good' instead of 'beneficial'.

Abstractness Paivio (1971) showed concrete words to be easier to understand than abstract ones. Qualities like ‘risk’, ‘optimism’, ‘security’ are inherently harder to process than objects or people (eg, ‘your home’, ‘your children’).

A word about jargon

Any group of people develops its own characteristic way of speaking: observational research shows that not only does language style vary by geographical region and social status, but also by gender and occupation. Members of the same profession or company learn to speak in their own specialised way.

When people complain about jargon it is usually about words that have leaked out from specialist contexts into general ones. It is usually carelessness, but in the worst cases it is evasive (politicians often use jargon in this way: see Bolinger (1980) for an old, but still entertaining and relevant summary of the workings of the ‘jargonauts’, as he calls them).

However, if you always insist on replacing jargon with common words (or, more often, phrases), you can end up with very long and complex sentences that are just as hard to process.

For example, the pensions industry has started to use the term ‘lifestyling’ to mean ‘the process of moving people’s money out of riskier investments into safer ones as they approach retirement age’. This means that, having explained the term properly, and occasionally revisiting that explanation, writers can call up that whole concept in a single word. Specialist terms can be used if they are properly introduced and explained, and if you remind readers of their meaning from time to time.

We are also aware of unpublished surveys of language preferences that show that some people find excessively plain English somewhat patronising, and expect official documents to have some degree of formality – so long as they can be understood.

What we look for

We look for vocabulary that the user is likely to have been exposed to – this may be a common sense judgement but it could be determined by checking the word frequency in a relevant corpus (a corpus is a collection of texts that are digitised for this kind of analysis). This needs to be judged in relation to the reader’s assumed level of expertise. We also watch for unfamiliar metaphors (common in financial reports – ‘market turbulence’ and ‘falling sentiment’

may not be understood). We check that any technical language either uses terms familiar to the audience, or provides layered explanations for those less expert (definitions, glossaries). We try to avoid talking pejoratively about ‘jargon’ – we just check that words are likely to be understood.

“What I took final exception to, was ‘to avoid skills fade’. You can’t just fool around with the language like that!”

1.3 Grammar and punctuation: conformity with the practice of good standard English

Grammar and punctuation can be emotive subjects, but rarely because they cause a failure in communication. If someone uses a double negative, or says ‘ain’t’ we understand exactly what they mean, but we might criticise them for using informal language in a formal setting.

Reasoning and evidence

Punctuation in particular is a hot topic, particularly since the publication of Lynne Truss’s best-selling book *Eats, Shoots, and Leaves* (2005). Luckily, in spite of her title, few people get shot because of misplaced commas (see Crystal 2006 for an antidote), and the main problem caused by poor punctuation or poor grammar is damage to reputation. It is an important signifier of literacy and competence, and poor punctuation is likely to attract vociferous complaints out of all proportion to any real damage to clarity.

There may be a generation gap here. Until the 1970s most schools taught grammar in quite a formal way, and students were taught to parse sentences (that is, to identify the structure of the sentence (for example, subject-verb-object), and the grammatical roles of its words (verbs, nouns, adjectives, etc). So it may be more important to use formal correct grammar if you know you are addressing older readers who may be more sensitive to it.

Grammar refers to rules for combining words into clauses and sentences. There are degrees of grammaticality which range from rigid rules that must be followed to stylistic preferences that act as genre markers (that is, they tell you how formal a document is, or what kind of assumption it makes about its audience).

An example of a rigid rule is the one that says that nouns and verbs need to agree whether they are singular or plural: ‘Send us one payments’ might leave you wondering what is actually meant. An example of a stylistic rule is the one that says that every sentence must include a verb – an action word. In formal English this may

be true, and we would find it very difficult to write or read many sentences without verbs. But creative writers, such as novelists or advertising copywriters, can be more flexible, and often include short sentences with no verb. Like this one. It gives a text a particular rhythm or emphasis.

Punctuation originally had a mainly prosodic role (that is, it marked pauses and intonation for reading aloud), and it is still used in this way in the language of advertising. In information text, the main role of punctuation is to make the division of clauses and sentences clear: to chunk language so it is easier to read and understand. Cohen et al (2001) found that punctuation in written language has a very similar effect to prosody (stress, pause, and intonation or ‘tune’) in spoken language. They also found that an absence of punctuation has little effect on comprehension, whereas deliberately wrong punctuation had a strong effect. Baron (2001) reviews the history of punctuation and theories about its function in language.

De Beaugrande (1984) identified a range of roles for punctuation, each designed to help overcome the linearity of writing through, for example, interpolation (parentheses), and parallelism (series of comma-separated items).

What we look for

We take grammar and punctuation to be an ‘entry condition’ and because we have mistakes are rare (and usually have little effect on communication) it is given a low weighting on our scorecard. We look for the use of standard English usage, but we are not pedantic about it, particularly where creative writing is concerned.

“It’s clear, absolutely simple, straightforward English, easy to understand. I hope all their communications are as easy to follow. If not I’ll just have to go and tell them that I thought this was wonderful and couldn’t they just...”

1.4 Readability: the ease with which the reader can follow the argument of the text

Readability refers to the ease with which language can be understood. It is affected by such things as the length and complexity of sentences, and the way they connect into longer narratives. Readability is different from legibility which refers to the ease with which we can physically read the words on paper or screen.

Reasoning and evidence

There is plenty of evidence that plain English makes documents easier and quicker to read. Many of the key research findings were established many years ago: plain English means using short sentences (Coleman 1962), concrete (rather than abstract) words (Paivio 1971), positive (rather than negative) constructions (Just & Clark 1973). This evidence mostly looks at speed of reading, and accuracy of recall.

The term ‘readability’ is often associated with the **readability formulas** that have become well known through their inclusion in the spelling check function of Microsoft Word. Because they are so available and apparently authoritative, it is worth taking a few moments to review how and whether they work.

Readability formulas have been around for many years, and one of the earliest developed, the Flesch Reading Ease formula, is still among the most commonly used. Rudolf Flesch (1948) established a model for formula development: identify the key variables that you think affect difficulty; produce versions of a text that alter those variables; test those versions on people who can be stratified in some easily communicated way (ie, school grade levels); then relate the count of those variables to the ability levels of the readers who succeeded in reading successfully.

Formulas are particularly attractive for institutions, since they can be seen as a cheap substitute for user testing, and they can also be used as an enforceable standard. Textbooks publishers often use readability scores as an acceptance criterion for manuscripts, and technical contracts sometimes use them when specifying performance levels of user guides.

The chief argument against readability formulas is that they do not fully represent the various causes of difficulty in text. They are easily manipulated, as writers can simply substitute short words for long ones, and cut sentences up, in order to achieve a better score.

In fact, what is so attractive about the idea of a formula is also at the heart of the main criticism: that a simple count of one or two variables can predict the outcome of a process as complex as reading and understanding.

George Klare (1984) reviews the history of readability formulas, and the arguments for and against. The attraction of formulas is that they are quite accurate in predicting reading difficulty, more so than many individuals. They correlate well with group scores (and they would have little credibility if they did not), but Klare reviews evidence that individual judgements are often unreliable.

Klare reviewed books on clear writing, and counted 156 different variables listed as important by experts. Many of these have been tried as constituents of readability formula, but as Klare remarks”

‘formula developers discovered rather quickly that many variables interrelate and, therefore, that continuing to add variables quickly runs into diminishing returns in predictive power... In general, two factors stand out, semantic difficulty and semantic difficulty.’ (page 714-715).

Semantic difficulty means difficulty with the meaning of words, and it is represented in Flesch’s formula by word length (shorter words being more common in everyday speech than longer ones, which are likely to be specialist or technical). In our benchmarking exercise, this is covered by the *Plain words* criterion.

Syntactic difficulty refers to the construction of sentences and clauses, and is represented in the Flesch formula by sentence length. The difficulty with longer sentences is partly due to the greater cognitive capacity that may be required to process them. Of course this depends on how well they are written and punctuated. It is particularly hard to read a long sentence in which you have reach the end before you can tie together the various sub-clauses.

A critical aspect of readability, but not covered by the main formulas, is **coherence**. This refers to the way in which ideas are connected across a text. Coherence is important both at a global level (across the whole text, or even across a set of texts) and at a local level (that is, in the way that ideas are connected within a paragraph).

Global coherence is covered in our benchmarking by the *Structure* criterion, so it is local coherence that this readability criterion mainly focuses on.

A good example is given by Anderson & Armbruster (1986), who quote a sample from a real textbook:

“In the evening, the light fades. Photosynthesis slows down. The amount of carbon dioxide in the air builds up again. This buildup of carbon dioxide makes the guard cells relax.”

This text has been written as short sentences, and achieves a creditable Flesch score of 69.5% (grade level 5.4). However, Anderson & Armbruster point out that the reader is left to infer the connections between these statements. Their rewrite makes these more explicit:

“What happens to these processes in the evening? The fading light of evening causes photosynthesis to slow down. Respiration, however, does not depend on light and thus continues to produce carbon dioxide. The carbon dioxide in the air spaces builds up again, which makes the guard cells relax. The relaxing of the guard cells closes the leaf openings. Consequently the leaf openings close in the evening as photosynthesis slows down.”

But the readability formula now thinks this is more, not less difficult: the grade level has gone up to 8.5, providing a caution against the use of formulas as part of the writing process, as well as a demonstration of the need for cohesion.

What we look for

We look for excessively long sentences, and for elements within sentences that can cause problems, such as deeply-embedded relative clause structures (that is, clauses containing words like ‘which’ and ‘who’ nested inside each other). We also look for well-connected arguments with each idea leading logically to the next. There should be no ‘so what’ moments, and our questions should be anticipated and answered.

2

Design criteria assess the visual impact of the document and the way its design influences usability.

“I don’t like that coloured bit, cause I can’t read it properly. Where it says what’s not covered, I can’t actually read it properly.”

“This is really tiny on the back, ridiculous, and it’s grey again! I mean why do they do it grey because it’s small as it is. It’s even worse.”

“It is simple, the writing is nice and bold.”

2.1 Legibility: use of legible fonts and text layout

Legibility relates to our ability to physically read the words on the page or screen. It is distinct from readability, which relates to the ease of understanding the meaning of the text.

Reasoning and evidence

If people physically can’t read your document, its content is irrelevant. Even if they can read it with effort, they are less likely to start, and there is evidence that less legible type not only slows readers down, but they read less accurately (Wright & Lickorish 1983).

Legibility is affected by a wide range of factors, including font design, line length, spacing and other things. Much of the research took place many years ago and the clearest review is still the one by Spencer (1968). It is worth noting that each factor (for example, boldness or line length) may produce quite small differences in legibility on its own, but in combination they can cause real problems.

One of the most commonly found legibility problems is an overlong line. Tinker (1963) recommends an optimal length that results in between 50-70 characters per line. Line length is, therefore, relative to type size, with smaller sizes needing narrower columns.

Research with people with normal or corrected sight shows that type as small as 8pt Times New Roman is acceptable (Hartley 1978), but that between 9 and 10pt is optimal for most purposes (Tinker 1963). The abbreviation ‘pt’ stands for ‘points’ – a system of measurement used for typography. It is worth noting that it relates to the total vertical space occupied by the letter, from bottom of the descender (the bottom part of letters such as ‘p’ or ‘y’) to the top of the ascender (the highest part of letters like ‘h’ or ‘d’). The relative size of the x-height (the main body of the letter) may vary with the design of the type, so expressions such as ‘10pt’ are quite approximate unless we specify a particular typeface.

Organisations who work on behalf of people with visual impairment make recommendations about legibility, and their requirements effectively call for special versions of documents (a 2mm x-height, corresponding to around 12pt in most typefaces is seen as the minimum for RNIB's less demanding 'clear type' standard, with at least 16pt required for large print). Research on legibility for people with visual impairment is reviewed by Russell-Minda et al (2007). We review some of this evidence in more detail in *Technical paper 10: Type size: interpreting the Clear Print standard*.

What we look for

We look for type that is too small, of an unusual or distorted design, printed with poor background contrast. Line length and interline spacing also contribute to legibility. We don't apply an absolute size limit and are more liberal in our judgements than the RNIB's recommendation of 12pt minimum for all type.

"I find usually pictures and instructions of how to put things together very very vague and you're not sure if you're doing the right thing or not. Pictures don't seem to bear any resemblance to what's in front of you, but that's a gem. It has a front view and back view, remarkably well labelled and useful arrows."

2.2 Graphic elements: use of page layout, tables, bullet lists, charts, diagrams, illustrations, etc

This criterion relates to the use of typography to make the structure of information clear, and the use of alternative to prose – pictures, charts, tables and lists.

Reasoning and evidence

The main purpose of typography and layout is to make the organisation of a document visually clear, so much of the discussion of the 'Structure' criterion is also relevant here.

But typography can help readers at the level of detailed explanations, as well as at the structural level of a document. Research confirms, for example, that typographic cues such as underlining, bold and colour help readers to understand and recall the information that is cued (Glynn & DiVesta 1979), although it must not be over-used.

A further strand of research has looked at graphic alternatives to prose such as a tables, lists and flowcharts. This is particularly relevant to financial documents, as it has been shown than conditional information is easier to understand when choices are shown graphically, or diagrammatically (Wright & Reid 1973; Felker et al 1981).

Layout has also been studied and found to affect the ability of readers to read strategically (Waller 1987), and to integrate arguments presented graphically – for example, to compare two competing viewpoints or options (Waller & Whalley 1987).

What we look for

We look for the appropriate use of typography to make hierarchies of information clear, and to distinguish between different kinds of information. Where tables are used, are they easy to read, with clear alignments? We look at the layout for the good use of space, lines and colour, and we look at the quality and appropriateness of illustrations, tables and charts – are they helpful, not misleading, placed well, and well designed?

2.3 Structure: quality of the document's organisation in relation to its function

This criterion looks at the organisation of the document or website – how easy it is to navigate, how easy it is to see how different parts are related, and how the part you are reading fits within the document or process as a whole.

Reasoning and evidence

This criterion, along with others in our benchmarking process, recognises that reading a document is not a passive linear activity in which language is interpreted and stored in memory. Communicating is not about injecting knowledge into the reader's bloodstream. It's more like providing a buffet from which they can choose: they have the job of selecting, eating and digesting.

Reading is an active process that involves the reader in using information strategically, in order to build an understanding of the message content, or to answer specific questions that they have. Effective readers vary the pace and order of their reading, which they treat as a problem-solving process (Thomas 1976). This active reading is only possible within a well-organised text, and preferably a visually organised one.

Research on reading strategy (the term used to describe what readers actually do, and why) is reviewed by Paris, Wasik & Turner (1991). The cognitive processes involved are known variously as *metacognition* (Brown, 1980; Baker & Brown, 1984) or *executive control processes* (Britton & Glynn 1987).

“Very clear step by step statement of what actions need to be taken.”

“It's quite confusing, what's not covered, and what is covered. It's all bitty and mixed up together somehow. It lists, miles and miles of it you see. And then it says please note that exclusion 15 and what is not covered by this policy. Blah blah blah. So then I have to go back and see what's that about. And then I couldn't find 15 for some time. Because it's not list 15 on the left hand side where I would have thought it would be, it's listed 15 in little letters here. I think it could be laidout more clearly. There's so much of it.”

Well-organised texts which are designed to support the full range of physical and mental processes that readers bring to them have been called ‘considerate texts’ (Armbruster 1984).

What we look for

We make a judgement about how useful and usable a document is, in relation to the processes it is expecting customers to engage in. These processes range from understanding the overall purpose of the document, to navigating within it, and understanding the meaning and significance of its content. It should be very clear from the visual organisation of a document how it is broken into sections, and how those sections relate.

“This would make you really want to read this because it’s more bolder writing and it’s got like certain parts and things like highlights and diagrams, so you would be interested to read it yes.”

“Professional, to the point, quality paper, minimal. Easy to store, good. Snobby quality of paper make feel professional and trust it. Nice positive language.”

2.4 Impression: attractiveness and approachability of the document’s overall appearance

This criterion is about how a document looks at first viewing. Do I want to read this? Do I want to read it now? How much attention shall I pay to it?

Reasoning and evidence

It has been found that interfaces that look attractive are also judged as easier to use, whether or not they actually are (Kurosu & Kashimura 1995). So attractiveness has a strong effect on initial uptake.

Research commissioned by Royal Mail (Harper & Shatwell 2002) showed that a typical family sorts mail on delivery into three categories, based on first impressions and the identity of the sender: read now, read later or discard without reading. ‘Read later’ mail may actually never be read, but perhaps filed in case it is important.

This read-now or read-later judgement is based largely on the first impression. A document is less likely to be read if it is seen as too difficult, too trivial, or as a sales document.

The first impression does not just affect whether/when a document is read, but the way in which it is read. Linguists who analyse documents see them as falling into different groups, known as genres or text types (Swales 1990). Each genre triggers strong expectations about how it will be organised, and how to read it – examples of common genres used by service brands or government departments are: letter, form, business terms or regulations, leaflet.

Genres are initially identified by readers at first glance, so layout, headings and format are all important influences on the kind of attention a document is likely to receive.

What we look for

We look for attractive documents with uncluttered but informative covers and early pages. We look for a clear writing style (particularly for headings), an orderly structure and clear, open layout. Each genre should work according to its own conventions – for example, letters should be short and to the point, without turning into sets of business terms; leaflets can be less formal, but should be graphically structured and inviting.

3 Relationship criteria assess how far the document establishes a relationship with its users.

“I hadn’t heard of this company before so I didn’t know what it was or read it until they sent me a second letter, when I thought it might be important.”

“If I hadn’t been reviewing my paperwork a little more, I would simply have chucked this thinking it was direct mail”

3.1 Who it’s from: is it clear who is communicating?

The perceived personality and authority of the sender of a document will affect how readers receive it. Most organisations have a defined brand, which can be invoked through a consistent use of the brand mark, colour, fonts and tone of voice.

Reasoning and evidence

The Royal Mail research cited earlier showed that people treat mail differently depending who it is from (Harper & Shatwell 2002). Mail from an organisation they do not have a relationship with is likely to be seen as junk mail, for example, and rejected.

If people can recognise the originator of a communication, it determines significantly the way in which the information will be viewed. Petty & Cacioppo (1983, 1984) proposed a communication model that suggests that readers are persuaded of the authority of a message partly through systematic analysis of its content, and partly by assessing the credibility of its source. They suggest that the source becomes more important in cases where the reader is unmotivated or lacks the skill to analyse the content systematically. A trusted brand or ‘speaker’ will create more of a tendency in readers to give credibility to the information. Readers are also likely to have an idea, however vague, of the sanctions that are in place (for example, regulation, journalistic scrutiny) that make it in the interests of well-known organisations to be truthful and to act in the interests of their customers – or at least, not to their detriment.

In the case of financial services, it is important to remember that people have very long relationships their providers – particularly in pensions, investments and life assurance. They may only hear from their provider once a year, and so changes in ownership, name or branding may take some time to filter through to customers. It may be necessary to remind customers of this, long after it has ceased to be a novelty in the organisation itself.

What we look for

The documents we look at in our benchmarking are mostly transactional and functional – they are sent to people as part of a process they have already engaged in as customers or enquirers (as distinct from unsolicited marketing material). So we look for very clear identifiers of the sender, and that there are unlikely to be any misunderstandings about its origins or status.

“If you don’t know what you’re doing, in big letters there’s a helpline thing here. It’s excellent.”

“This doesn’t have a name of who to contact if you have any questions – I don’t think that’s very good.”

3.2 Contact: whether there are clear contact points

Clear communications channels are obviously important for processes to be efficient, and they are also a matter of courtesy to our customers who want to be able to use the channel they prefer.

Reasoning and evidence

All communication is part of a dialogue, and success depends on this working both ways.

Research on user guides (a different but equally complex area of communication) has shown that most people strongly prefer to ask a human expert rather than to try to understand a complex document (Wright 1981; Wright, Creighton & Threlfall 1982). For most people, asking someone is the preferred way to overcome difficulty.

Clear contact channels also relate to a key concept in usability research: *forgiveness* (reviewed by Lidwell, Holden & Butler 1983). No interface can be expected to work for every user, but successful ones are forgiving of error – offering an alternative to failure. An unforgiving communication is one that, if it fails, offers the user no alternative. In this case, the forgiving alternative is to call, which means that the contact details are crucial.

What we look for

We look for clear contact details provided in a prominent place. If premium phone lines are used (for example, 0845 numbers), the cost should be made clear, and opening hours should be shown. Customers should also be given the option of email or postal mail.

Alternative routes for people with special needs should also be clearly shown.

Other addresses or apparent contact details (such as the company registration) should not distract from the main contact channels.

“Compared to the other financial example which was obviously more ‘high powered’, this type of financial information is ‘for the peasants’. It explains everything, and is perhaps a bit patronising, but I like that because I don’t really know what all of it means.”

3.3 Audience fit: appropriateness to the knowledge and skills of the users

This criterion is related to *Relevance*, and also to *Plain words*, but focuses on the level of conceptual difficulty that is appropriate for the audience.

Reasoning and evidence

Surveys show that a large number of people in the UK fall below a satisfactory level of functional literacy (DfES 2003): this term describes the ability to use documents for problem-solving tasks, in addition to the simple deciphering of words and sentences (this is known as prose literacy). Functional literacy also includes numeracy, and this is a particular problem for many people. For example, a great many people have difficulty with the concept of percentages – a fundamental concept in the financial services industry.

In recent years, specific literacies have been identified and researched, including financial literacy (Atkinson et al 2007). The Financial Inclusion Taskforce has been established by the government to address this particular problem. Data from this research can be used to alert writers to specific issues that customers are likely to have difficulty understanding.

Difficulty will result from any mismatch between the reader and the text. This may be at many levels: in the content (for example, complex decisions involving risk assessments), in the expression (difficult words and complex explanations) or in the reader’s motivation.

Research also shows that readers’ interest in a topic can affect performance (Baldwin et al 1985, Stevens 1980). This is bad news for writers about financial products, as a separate strand of research known as behavioural economics has identified psychological factors related to time-distant outcomes of decisions (Lieberman et al 2002). This effectively confirms what is generally known: that younger people have great difficulty in thinking through pension options that relate to a distant future. This work suggests that overcoming difficulty may involve motivating customers as well as simplifying the document.

What we look for

We look for a level of conceptual and language difficulty that is appropriate for the audience, or for a clear layering of information, or alternative reading paths to help different users get the most

from a document or website. Communications aimed at financial advisers, for example, obviously need a different level of explanation from communications aimed at their non-expert customers.

“Harsh and cold messaging - could be written in a more positive way.”

“The ‘simply do x’ language is irritating – what if I don’t find doing it simple? Then you are saying I’m stupid!”

3.4 Tone: matching the style and language to the context

This criterion looks at the overall style of the language you use. Although it could belong in our language section, this covers those aspects of language that create relationships.

Reasoning and evidence

This aspect of language is studied by discourse linguists or sociolinguists (Brown & Yule 1983; Hudson 1996). It is typically descriptive, rather than empirical (that is, it is not usually based on experiments), and offers useful insight about why plain English is so hard to achieve for many organisations.

It is worth noting that the concept of ‘everyday’ English is not straightforward and should not be taken too literally. Many people’s everyday English is informal and non-standard (in the sense that it represents a local or group dialect, rather than the standardised grammar heard on the BBC news). Brands that use truly informal English for administrative documents risk losing credibility with groups whose own informal English is very different (compare the tone of voice used by Virgin and Saga, for example).

One area where the research is not clear-cut is the comprehension of metaphorical language, which is very prevalent in descriptions of investment performance (for example, metaphors of war, mechanics, weather and health are common: confidence might be ‘battered’, investment strategies ‘aggressive’ or ‘defensive’; growth is ‘generated’ by ‘dynamo’ economies, and crises ‘sparked’; the outlook might be ‘sunny’ or ‘stormy’; sectors might be ‘ailing’ or ‘healthy’.) Although metaphor might be thought to be an indirect way of speaking, and therefore unreliable, research reviewed by Gibbs & Steen (1999) shows this not to be the case. We have not found research specifically related to financial products, though, nor to any other domains of interest to our member organisations.

What we look for

We look for an appropriate tone of voice for the topic, the audience and the organisation's brand. For most of our member organisations, the starting point is clear, everyday English, with some variation where brands want to be particularly distinctive (for example, through informality) or where they address audiences which may include many people with poor education, or who speak English as a second language. Documents might lose marks here if they appear to trivialise important issues, if they appear not to care about how they are understood, or are excessively self-regarding or promotional in inappropriate places.

4 **Content criteria** assess how the content and the way it is organised deliver the document's purpose.

“Not quite sure why I got this – maybe because last year I qualified for a disability living allowance.”

“Relevance of the second page of ‘additional conditions’ not clear as the account has matured”

4.1 **Relevance: how relevant the content is to the recipient**

We are all more interested in information we see as relevant to us. Because it fits with our needs and expectations, we attend to it and draw conclusions from it.

Reasoning and evidence

Relevance is a key question that affects the customer on a number of different levels: it is not only part of the initial impression that motivates them to read on, but it also affects the difficulty of a document. This is because it reflects the degree to which a document is personalised.

The most personalised communication is a phone call, or face-to-face meeting. So long as both parties feel free to raise topics or ask questions, relevance is guaranteed. Almost any other channel involves the organisation guessing what the customer needs, and the customer guessing what the organisation is trying to tell them. Faced with information that is highly generalised or of questionable relevance, the customer may find no existing mental schemata within which to interpret it. They have to invent a possible world in which it might be relevant to them – quite apart from the effort required, they risk getting it wrong.

Apart from a conversation, the next most personalised channel is an individually written communication, followed by a standard document containing personal data and selective messages, then a standard letter sent to a particular customer segment, and finally an indiscriminate standard document.

As we move down this list, the customer is being asked to work harder to find content that applies to their situation, and to weed out messages that should be discarded. This obviously carries risks that they may become confused or demotivated, or even take a completely wrong meaning from the document.

Information that is not highly personalised normally includes conditional information signalled by expressions such as ‘if’, ‘or’, ‘except’, ‘in some cases’, or ‘customers with [product] should’. Or the conditionality may be implicit, expressed through a choice of section headings to select from.

Research has shown conditional information to be particularly difficult to understand (Neisser & Weene 1962, Holland & Rose 1980), and more so where negative conditions are involved (eg, ‘if you do not have...’) (Just & Clark 1973, Sherman 1976; Wright & Hull 1986). Conditional information involves an increased cognitive processing load, which is often made worse in poor documents by other sources of complexity (such as difficult words or long sentences, for example).

Further evidence of the need for relevance to be clearly signalled can be found in the research that underlies the measurement of *document literacy*. This term describes the ability to use documents for functional purposes (Kirsch 2001), as distinct from *prose literacy*, which is the ability to decode words and sentences. One of the key components of document literacy is the ability to find the correct answer to a question from among a series of incorrect, but plausible answers, known as distractors (see a good explanation by Evetts & Gauthier 2005).

What we look for

It is not always possible to guarantee a message will be seen as relevant, but there are things that help. We look for indications of the expected audience (‘a guide for new parents’, ‘employing someone for the first time?’), layered information that makes relevance clear for particular audiences, clear headings, diagrams and flow charts.

“This is difficult and I wish they'd put the name of the insurance policy and company as well as the policy ref. It would have meant more immediately if I'd known it related to the Household and contents policy”

4.2 Subject: whether it is clear what the communication is about

This criterion is a critical one, as it addresses the kind of major misunderstandings that can happen if participants in a conversation are at cross-purposes.

Reasoning and evidence

At the start of this paper we introduced the concept of schemata. A problem for specialist documents is that people may not be familiar with the relevant concepts and processes and do not have well-developed mental models to apply. So when you write a letter, you need to do more than quote an account number or even a product brand name: you also need to identify the subject in the customer's own terms, and your purpose in writing the document. For example: 'Changes to your pension', 'Your benefit is going up', 'You are at the age you told us you want to retire'.

Establishing the subject may also involve explaining the background and setting up a framework that explains the topic of the document. This kind of pre-summary has been called an 'advance organiser' (Ausubel 1963). Research has confirmed the effectiveness of advance organisers and summaries for orienting readers to topics (Hartley & Davies 1976).

What we look for

This criterion is closely related to *Relevance*, but focuses on how the topic and purpose of the communication are announced. We look at the titles, subtitles and prominent summaries, and how clearly they relate to the world of the customer.

“We are not sure what to do with it – where is the form to complete?”

“Very clear step by step statement of what actions need to be taken.”

4.3 Action: clarity about what action is required of the user

Most functional communications potentially lead to actions – either a necessary action – sending something in, paying or phoning – or a potential action such as making an insurance claim, reporting a change of circumstances, or investing some money.

Reasoning and evidence

Anecdotal evidence from many organisations tells us that mistakes are extremely common in form-filling and bill-paying. Our experience as designers tells us that any form that has to be filled in by the entire population (with the full spectrum of literacy levels) will have an error rate of at least 10-15%, rising to much higher figures in the case of longer or more complex forms. People pay the wrong amount, send cheques with no identifying information, or send things to the wrong address. In the mid-1990s the UK Passport Agency reduced errors in passport applications by around 6% simply by providing an addressed envelope for customers to use. The envelope had the correct address, and a checklist of what to enclose.

Even in communications that are mostly informative, the *Action* criterion represents the customer’s need to understand the overall process that this communication is part of. We have already discussed the concept of mental models or schemata, and there is some additional evidence that step-by-step structures are effective for communicating technical concepts. (Novick & Morse 2000). This means that a clear Next Step is the last piece of information that helps readers complete this piece of communication – whether it’s providing the answer to a question, or taking in some information. Without understanding of the Next Step, the communication will be incomplete. This means that the dialogue between the organisation and the customer has broken down.

What we look for

We look for clarity of instructions and communication routes (also covered by our *Contact* criterion). Are amounts clear, are conditions clear, do people know where to sign? This criterion becomes critical with forms, which need to relate response boxes clearly to questions, and bills, which need to be very clear about amounts and dates.

“The only minor criticism it took me a bit of time to find out what to do if I didn’t want to reinvest, because they want you to reinvest, everything about reinvesting is ‘what to do next’, how to reinvest etc all clear. And then in very small, well ordinary print here, if we don’t receive your reinvestment form by the x date, or you would rather not invest, we will automatically pay it out. Thank-you, that’s all I need to know.”

“I have a mortgage with (this company) and they continually send me letters about loans/insurance etc even though I’ve told them I’m not interested. Ethically, I think it’s wrong for them to try to encourage people to take out loans in this way. I’ve lost respect for the company.”

4.4 Alignment: compliance with the organisation’s intended aims and values

This criterion considers whether the effect of the document is likely to be in line with the intention of the organisation that produced it. This takes into account the fact that people’s decisions are influenced by non-rational factors, and that messages need to be designed to fit with user expectations.

Reasoning and evidence

Just because a document is clear and understandable, it does not mean it will have the intended effect. We introduced this criterion to allow us to evaluate the role of the document in the conversation between an organisation and its customers. It asks whether a document is likely to have its intended effect, bearing in mind the customer’s information needs and expectations.

The alignment criterion is underpinned by two main sources of reasoning: firstly, theories and research evidence from behavioural economics that address the way in which people make decisions; and secondly, the notion that documents are part of a conversation between the organisation and its customers, in which each has an expected role to play.

Most of the documents we benchmark exist to support, directly or indirectly, choices that people are being asked to make. The theory of choice has been grounded in economics. A key assumption by classical economics has been that individuals take decisions so as to maximise their own welfare. In doing so they weigh the costs of actions against their benefits, and where costs or benefits are uncertain, they discount them in a consistent way.

However, recent years have seen the emergence of behavioural economics – the use of psychological methods and theories to explore decision making. This has increasingly shown that in reality people do not make decisions in the way classical economists assumed, especially in conditions of uncertainty. For example, it turns out that we value possible gains less than possible losses, even when the amounts and probabilities are the same. (Kahneman & Tversky 1979). We use heuristics – rules of thumb – rather than calculation (Gigerenzer & Goldstein 1996). And when faced with too many options we are less likely to make any kind of choice, and may give up and do nothing (Iyengar & Lepper 2000). The Financial Services Authority recently commissioned a good summary of the

relevance of behavioural economics to financial decision-making (de Meza, Irlenbusch & Reyniers 2008)).

The result of these factors is that it may not be possible for the authors of functional documents to avoid biasing the customers' choices in one way or another. That leads to the argument that if we can't help biasing customers' choices, we should at least bias them in their own best interests (an argument widely disseminated through the best-seller *Nudge* by Thaler & Sunstein 2008). The difficulty for the Simplification Centre is that making independent assumptions about what decisions customers ought to make in their own interests is fraught with difficulty. We have no source of objective evidence beyond customers' own decisions, and we are unlikely to be able to analyse those in sufficient depth to reach soundly based conclusions. Anything else would require us to make our own value assumptions which may be different from those of organisations or their customers.

We therefore concluded that it would be more helpful if our benchmarking were to comment on the alignment between the likely effects of the documents and the values and objectives which the originating organisation declares them to perform. Studies such as Meyer & Schwager (2007) – as well as ubiquitous anecdotal evidence – show how frequently there is a divergence between an organisation's corporate aspirations for the quality of its service and the experience at the touchpoints – that is, where the customer or user comes into direct contact with the organisation. Kaplan & Norton (1992 and 1996) emphasised the key role of relevant information flows in measuring the alignment of service delivery with management's strategic aims.

This brings us to our second source of reasoning for this criterion. By defining a personality for their brands, and using a house style and corporate identity to ensure consistency in their communications, organisations strive to appear to a customer as if they are an individual. But this sets up high expectations that they will therefore have the same memory of events, commitment to promises and good manners that an individual normally has. Every encounter with a brand is received equally on this basis – whether it is a carefully crafted marketing letter or a long wait in a call centre queue.

Waller & Delin (2003) use the work of the philosopher H.P. Grice to develop the concept of cooperative brands. His Cooperative Principle develops the idea that people in conversation are cooperating in a mutual effort to communicate – he defines four maxims – principles

Thanks to a number of colleagues who commented on or contributed to this report, including Judy Delin, Martin Evans, John Field and others I may have forgotten about, as an early draft was circulated several years ago. They can take credit for any good things you might find here, but are not responsible for the faults you will find.

that each participant is entitled to assume the other is following, in order to make sensible inferences. The principles, paraphrased, are:

- Relation: Be relevant.
- Manner: Be as clear and brief as possible.
- Quality: Tell the truth, as far as you know it.
- Quantity: Give the right amount of information (neither withholding anything that is relevant, nor swamping with unnecessary detail).

It is the customer's natural assumption of cooperation that leads them to believe that they would not be offered a product they do not need, and that no important terms of business are being hidden. Recent rulings by regulators about the mis-selling of financial products would seem to confirm it is reasonable for them to make such an assumption.

These principles are also covered among our other benchmarking criteria, but they reappear here to enable us to take account of their effect on the alignment of communications with the stated intentions and brand personality of the organisation.

What we look for

We ask organisations whose documents we benchmark what they want it to achieve. We look for factors such as the way questions or propositions are framed, the number of choices people have to make and the way risks are described.

References

- Anderson, T., H. & Armbruster, B. B. (1986). Readable textbooks. In J. Orasanu (Ed.), *Reading comprehension: from research to practice* (pp. 151-162). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Armbruster, B.B. (1984). The problem of 'inconsiderate texts'. In G.G. Duffy, L.R. Roehler, & J. Mason (Eds.), *Theoretical issues in reading comprehension* (202-217). White Plains, NY: Longman.
- Atkinson, A., S. McKay, et al. (2007). Levels of financial capability in the UK. *Public Money & Management*, 27, 29-36.
- Ausubel, D. (1963). *The Psychology of Meaningful Verbal Learning*. Orlando, Florida: Grune & Stratton.
- Baker, L., & Brown, A.L. (1984). Metacognitive skills in reading. In P.D. Pearson (Ed.), *Handbook of reading research* (353-394). New York: Longman.
- Baldwin, R. S., Peleg-Bruckner, Z., & McClintock, A. H. (1985). Effects of topic interest and prior knowledge on reading comprehension. *Reading Research Quarterly*, 20, 497-504.
- Baron, N.S. (2001). Commas and canaries: the role of punctuation in speech and writing', *Language Sciences*, 23: 15-67
- Bolinger, D. (1980). *Language: the loaded weapon*. London and New York: Longman.
- Britton, B., & Glynn S. (eds). (1987). *Executive control processes in reading*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brown, A.L. (1980). Metacognitive development and reading. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brown, G., & Yule, G. (1983). *Discourse Analysis*. Cambridge: Cambridge University Press.
- Cohen, H., Douaire, J. & Elsabbagh, M. (2001). The role of prosody in discourse processing, *Brain and Cognition*, 46, 73-82.
- Coleman, E.B., 1962. Improving comprehensibility by shortening sentences. *Journal of Applied Psychology*, 46, 131-134.
- Crystal, D. (2006). *The fight for English: How Language Pundits Ate, Shot, and Left* Oxford University Press.
- De Beaugrande, R., (1984). *Text production: Towards a science of composition*. Norwood, NJ: Ablex.
- De Meza, D., Irlenbusch, B., Reyniers, D. (2008) Financial capability: a behavioural economics perspective. Consumer Research paper 69. London: Financial Services Authority
- DfES. (2003). *The Skills for Life Survey: a national needs and impact survey of literacy, numeracy and ICT skills*. Research Brief RB 490, London: Department for Education and Skills.
- Evetts, J., & Gauthier, M. (2005). *Literacy Task Assessment Guide*. National Literacy Secretariat, Canada.
- Felker, D.B., Pickering, F., Charrow, V.R., Holland, V.M., & Redish, J.C. (1981). *Guidelines for Document Designers*. Washington DC: American Institutes for Research.
- Ferreira, F. (2003) The misinterpretation of noncanonical sentences. *Cognitive Psychology* 47, 164-203.
- Flesch, R.F., (1948) A new readability yardstick. *Journal of Applied Psychology*. 32: 221-233.

- Gibbs, R. and Steen, G. (eds) (1999) *Metaphor in Cognitive Linguistics*. Amsterdam: Benjamins.
- Gibson, J.J. (1977). The Theory of Affordances. In *Perceiving, Acting, and Knowing*, R. Shaw and J. Bransford (eds). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gigerenzer, G., & Goldstein, D. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review*, 103(4), 650 – 669
- Glynn, S., & DiVesta, F. (1979). The control of prose processing via instructional and typographic cues. *Journal of Educational Psychology*, 71, 595-603.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Sadock (Eds.), *Syntax and semantics III: Speech acts*. New York: Academic Press.
- Harper, R & Shatwell, B. (2002). Paper-mail in the Home of the 21 st Century: An analysis of the future of paper-mail and implications for the design of electronic alternatives, *Journal of Interactive Marketing*, April/June, 3, 311-323.
- Hartley, J. (1978). *Designing instructional text*. NY: Nichols Publishing Company.
- Hartley, J., & Davies, I. (1976). Preinstructional strategies: the role of pretests, behavioural objectives, overviews and advance organizers. *Review of Educational Research*. 46, 239-265.
- Holland, V.M., & Rose, A.M. (1980). *Understanding instructions with complex conditions* (Document Design Project Technical Report No 5). Washington DC: American Institutes for Research.
- Horgan, Dianne (1976). The development of the full passive. *Journal of Child Language*, 5, 65-80.
- Hudson, R.A. (1996). *Sociolinguistics*. Cambridge University Press.
- Hupet, M., & LeBouedec, B. (1975). Definiteness and voice in the interpretation of active and passive sentences. *Quarterly Journal of Experimental Psychology*, 27, 323-330.
- Iyengar, S., & Lepper, M. (2000) When choice is demotivating: can one desire too much of a good thing? *Journal of Personality and Social Psychology* 76, 995-1006.
- Just, M., & Clark, H.H., (1973). Drawing inferences from the presuppositions and implications of affirmative and negative sentences. *Memory & Cognition*, 4, 318-322.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, 47(2), 263-292
- Kaplan, R., & Norton, D. (1992) The balanced scorecard – measures that drive performance. *Harvard Business Review*. 70(1), 71-79.
- Kaplan, R., & Norton, D. (1996). *The balanced scorecard: translating strategy into action*. Harvard Business School Press
- Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kintsch, W., & van Dijk T.A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363-94.
- Kirsch, I. (2001). *The international adult literacy survey (IALS): Understanding what was measured* (ETS Research Report RR-01-25). Princeton, NJ: Educational Testing Service.
- Klare, G. R. (1984). Readability. In D. P. Pearson, R. Barr, M. L. Kamil & P. Mosenthal (Eds.), *Handbook of Reading Research* (pp. 681-744). Mahwah, NJ: Lawrence Erlbaum Associates.
- Kurosu, M., & Kashimura, K. (1995). Apparent usability vs inherent usability: experimental analysis on the determinants of apparent usability, *CHI '95 Conference Companion*, 292-293.

- Lieberman, N., Sagristano, M.D. et al. (2002). "The effect of temporal distance on level of mental construal." *Journal of Experimental Social Psychology* 38(6): 523-534.
- Lidwell, W., Holden, K. & Butler, J. (1983). *Universal Principles of Design*, Gloucester, MA: Rockport.
- McNamara, T.P., Miller, D.L., and Bransford, J.D. (1984). Mental models and reading comprehension. In P.D. Pearson (Ed.), *Handbook of reading research* (490-511). New York: Longman.
- McVee, M.B., Dunsmore, K., & Gavelek, J.R. (2005). Schema Theory Revisited. *Review of Educational Research*, 75, 531-566.
- Meyer, C., & Schwager, A. (2007). Understanding customer experience. *Harvard Business Review*, 85 (2), 116-126.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97
- Nagy, W.E. and Herman, P.A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. McKeown and M. Curtis (Eds.), *The Nature of Vocabulary Acquisition*, (19-35). Hillsdale, NJ: Erlbaum Associates.
- Neisser, U., & Weene, P. (1962). Hierarchies in concept attainment. *Journal of Experimental Psychology*, 61, 451-460.
- Novick, L. R. , & Morse, D. L. (2000). Folding a fish, making a mushroom: the role of diagrams in executing assembly procedures. *Memory & Cognition* 28, 1242-1256.
- Paris, S.G., Wasik, B.A., & Turner, J.C. (1991). The development of strategic readers. In R. Barr, M.L. Kamil, P. Mosenthal, & P.D. Pearson (Eds.), *Handbook of Reading Research* (Vol. 2, 609-640). New York: Longman.
- Paivio, A., (1971). *Imagery & Verbal Processes*. New York: Holt.
- Petros, T. V., Bentz, B., Hammes, K., & Zehr, D. H. (1990). The components of text that influence reading times and recall in skilled and less skilled college readers. *Discourse Processes*, 13, 387-400.
- Petty, R. E., & Cacioppo, J. T. (1983). Central and peripheral routes to persuasion: Application to advertising. In L. Percy & A. Woodside (eds). *Advertising and consumer psychology*. Lexington: D. C. Heath.
- Petty, R. E., & Cacioppo, J. T. (1984). Source factors and the elaboration likelihood model of persuasion. *Advances in Consumer Research*. 11, 668-672
- Russell-Minda E., et al (2007), The legibility of typefaces for readers with low vision: a research review, *Journal of Visual Impairment & Blindness*, July 2007.
- Sherman, M. (1976). Adjectival negation and the comprehension of multiple negated sentences. *Journal of Verbal Learning and Verbal Behavior*, 12, 21-31.
- Slobin, D.I. (1966). Grammatical transformations and sentence comprehension in childhood and adulthood. *Journal of Verbal Learning and Verbal Behavior*, 5, 219-227.
- Spencer, H. (1968). *The visible word*. London: Lund Humphries.
- Stevens, K. (1980). The effect of topic interest on the reading comprehension of higher ability students. *Journal of Educational Research*, 73, 365-368.
- Swales, J.M. (1990). *Genre Analysis*. Cambridge: Cambridge University Press
- Thaler, R., & Sunstein, C. (2008). *Nudge: improving decisions about health, wealth and happiness*. Yale University Press
- Thomas, L. (1976). *The self-organized learner and the printed word*. Uxbridge, England: Brunel University Centre for the Study of Human Learning.
- Tinker, M. (1963). *Bases for effective reading*, University of Minnesota Press.
- Truss, L. (2005). *Eats, Shoots & Leaves: The Zero Tolerance Approach to Punctuation* (2005) Profile Books Ltd.

- Waller, R., & Whalley, P. (1987). Graphically organised prose. In E. De Corte, H. Lodewijks, R. Parmentier and P. Span (eds). *Learning and instruction*. Oxford: Pergamon Press, 369-381
- Waller, R. (1987). Typography and reading strategy. In B. Britton & S. Glynn (eds), *Executive control processes in reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Waller R., & Delin J. (2003) Cooperative brands: the importance of customer information for service brands, *Design Management Journal*, 14 (4), 63-69.
- Wheeler, D.D. (1970). Processes in word recognition. *Cognitive Psychology*, 1, 59-85.
- Wright, P. (1981). The instructions clearly state... can't people read? *Applied Ergonomics*, 12, 131-142.
- Wright, P., & Hull, A.J. (1986). Answering questions about negative conditionals. *Memory and Language*, 25, 691-709.
- Wright, P. and Lickorish, A. (1983) Proofreading texts on screen and paper. *Behaviour and Information Technology*, 2, 227-235.
- Wright, P., & Reid, F. (1973). Written information: some alternatives to prose for expressing the outcomes of complex contingencies. *Journal of Applied Psychology*, 73, 160-166.
- Wright, P., Creighton, P., & Threlfall, S.M. (1982). Some factors determining when instructions will be read. *Ergonomics*, 25, 225-237.