

Research behaviour in the new information landscape

Ian Rowlands
CIBER, University College London

UHMLG Spring Forum 2010
Woburn House Conference Centre

Reading the tea leaves

Deep log analysis

86.42.128.213 - - [02/Aug/2009:23:56:31 +0100] "GET /image/google-sm.gif HTTP/1.1" 200 654 "<http://www.slais.ucl.ac.uk/>" "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.0; Trident/4.0; GTB6; SLCC1; .NET CLR 2.0.50727; Media Center PC 5.0; .NET CLR 1.1.4322; .NET CLR 3.5.30729; .NET CLR 3.0.30618)"

86.42.128.213 - - [02/Aug/2009:23:56:31 +0100] "GET /image/xhtml1.0.png HTTP/1.1" 200 929 "<http://www.slais.ucl.ac.uk/>" "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.0; Trident/4.0; GTB6; SLCC1; .NET CLR 2.0.50727; Media Center PC 5.0; .NET CLR 1.1.4322; .NET CLR 3.5.30729; .NET CLR 3.0.30618)"

86.42.128.213 - - [02/Aug/2009:23:56:31 +0100] "GET /image/css.png HTTP/1.1" 200 918 "<http://www.slais.ucl.ac.uk/>" "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.0; Trident/4.0; GTB6; SLCC1; .NET CLR 2.0.50727; Media Center PC 5.0; .NET CLR 1.1.4322; .NET CLR 3.5.30729; .NET CLR 3.0.30618)"

Reading the tea leaves

Deep log analysis

86.42.128.213 - - [02/Aug/2009:23:56:31 +0100] "GET /image/google-sm.gif HTTP/1.1" 200 654 "<http://www.slais.ucl.ac.uk/>" "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.0; Trident/4.0; GTB6; SLCC1; .NET CLR 2.0.50727; Media Center PC 5.0; .NET CLR 1.1.4322; .NET CLR 3.5.30729; .NET CLR 3.0.30618)"

86.42.128.213 - - [02/Aug/2009:23:56:31 +0100] "GET /image/xhtml1.0.png HTTP/1.1" 200 929 "<http://www.slais.ucl.ac.uk/>" "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.0; Trident/4.0; GTB6; SLCC1; .NET CLR 2.0.50727; Media Center PC 5.0; .NET CLR 1.1.4322; .NET CLR 3.5.30729; .NET CLR 3.0.30618)"

86.42.128.213 - - [02/Aug/2009:23:56:31 +0100] "GET /image/css.png HTTP/1.1" 200 918 "<http://www.slais.ucl.ac.uk/>" "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.0; Trident/4.0; GTB6; SLCC1; .NET CLR 2.0.50727; Media Center PC 5.0; .NET CLR 1.1.4322; .NET CLR 3.5.30729; .NET CLR 3.0.30618)"

Strengths of DLA	Weaknesses of DLA
Comprehensive data, no sampling issues.	
What people actually did, not what they remember or invent	
Grounded theory, no prior assumptions	

Reading the tea leaves

Deep log analysis

86.42.128.213 - - [02/Aug/2009:23:56:31 +0100] "GET /image/google-sm.gif HTTP/1.1" 200 654 "<http://www.slais.ucl.ac.uk/>" "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.0; Trident/4.0; GTB6; SLCC1; .NET CLR 2.0.50727; Media Center PC 5.0; .NET CLR 1.1.4322; .NET CLR 3.5.30729; .NET CLR 3.0.30618)"

86.42.128.213 - - [02/Aug/2009:23:56:31 +0100] "GET /image/xhtml1.0.png HTTP/1.1" 200 929 "<http://www.slais.ucl.ac.uk/>" "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.0; Trident/4.0; GTB6; SLCC1; .NET CLR 2.0.50727; Media Center PC 5.0; .NET CLR 1.1.4322; .NET CLR 3.5.30729; .NET CLR 3.0.30618)"

86.42.128.213 - - [02/Aug/2009:23:56:31 +0100] "GET /image/css.png HTTP/1.1" 200 918 "<http://www.slais.ucl.ac.uk/>" "Mozilla/4.0 (compatible; MSIE 8.0; Windows NT 6.0; Trident/4.0; GTB6; SLCC1; .NET CLR 2.0.50727; Media Center PC 5.0; .NET CLR 1.1.4322; .NET CLR 3.5.30729; .NET CLR 3.0.30618)"

Strengths of DLA	Weaknesses of DLA
Comprehensive data, no sampling issues.	Little or no contextual information about task or motivation.
What people actually did, not what they remember or invent	Unit of analysis is usually the session, not the individual.
Grounded theory, no prior assumptions	What does it mean? (Value judgements inappropriate.)

What the tea leaves tell us

Deep log analysis

horizontal information seeking: skimming, viewing 1-2 pages of an online source, probably never returning

navigation: extended time navigating a site rather than viewing content

power browsing: very short dwell times, rapid clicking

squirreling behaviour: squirrelling away material, either by downloading it, bookmarking it, or cutting and pasting

checking: establishing the reliability of information by rapid cross-checking across multiple sites

In densely printed pages of text, reading is linear and strictly coded. Such texts must be read the way they are designed to be read: from left to right and from top to bottom, line by line. Any other form of reading (skipping, looking at the last page to see how the plot will be resolved) is a form of cheating and produces a slight sense of guilt in the reader.

Kress and Van Leeuwen, 2006: 204.

I am not thinking the way I used to think. I can feel it most strongly when I am reading. Immersing myself in a book or lengthy article used to be easy ... deep reading that used to come naturally has become a struggle. Once I was a scuba diver in a sea of words. Now I zip along the surface like a guy on a jet-ski.

Nicholas Carr, 2008:2.

the Web Behaviour Test

Discover your inner animal

We've identified 8 'species' of web user
What sort of web animal are you?

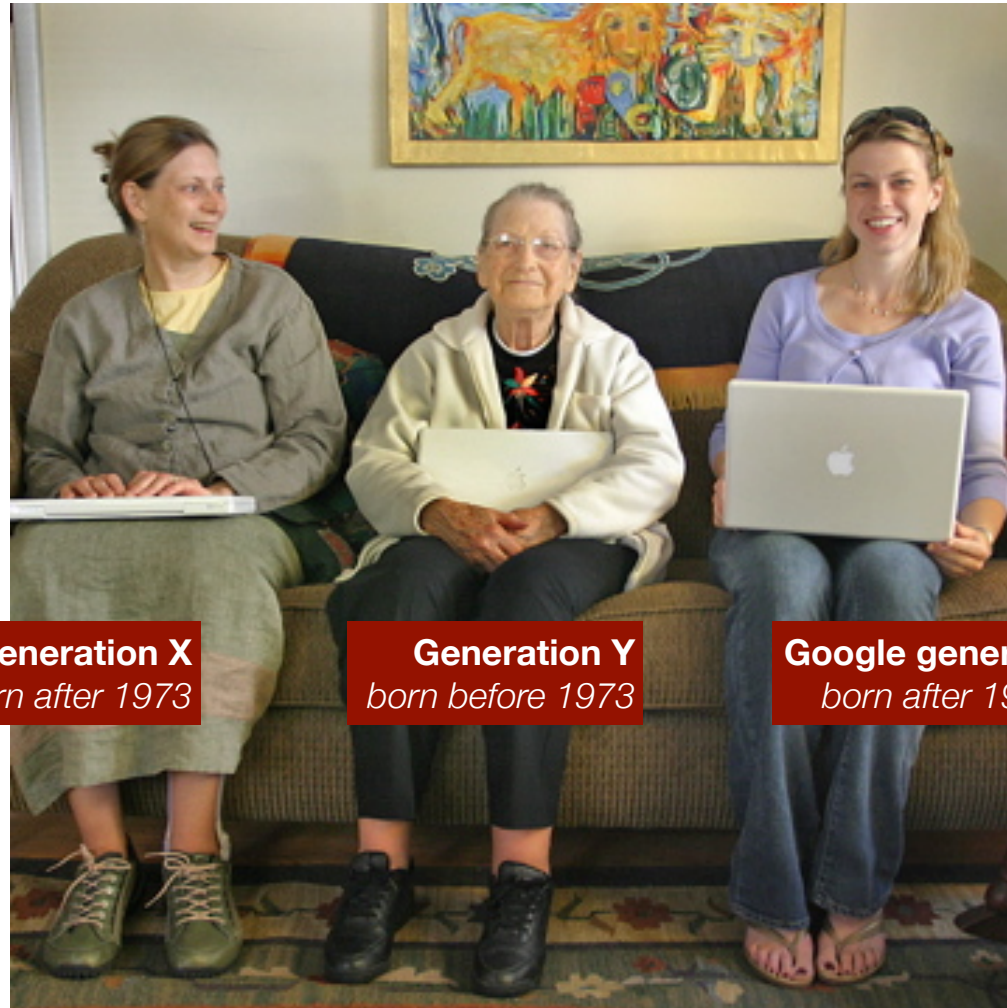
- Take part in real science
- Takes just 20 mins

[Start the Web Behaviour Test](#)

You will need to [sign in](#) or [register](#)

Age-related differences in information-seeking

BBC Digital Revolution experiment pilot phase



Generation X
born after 1973

Generation Y
born before 1973

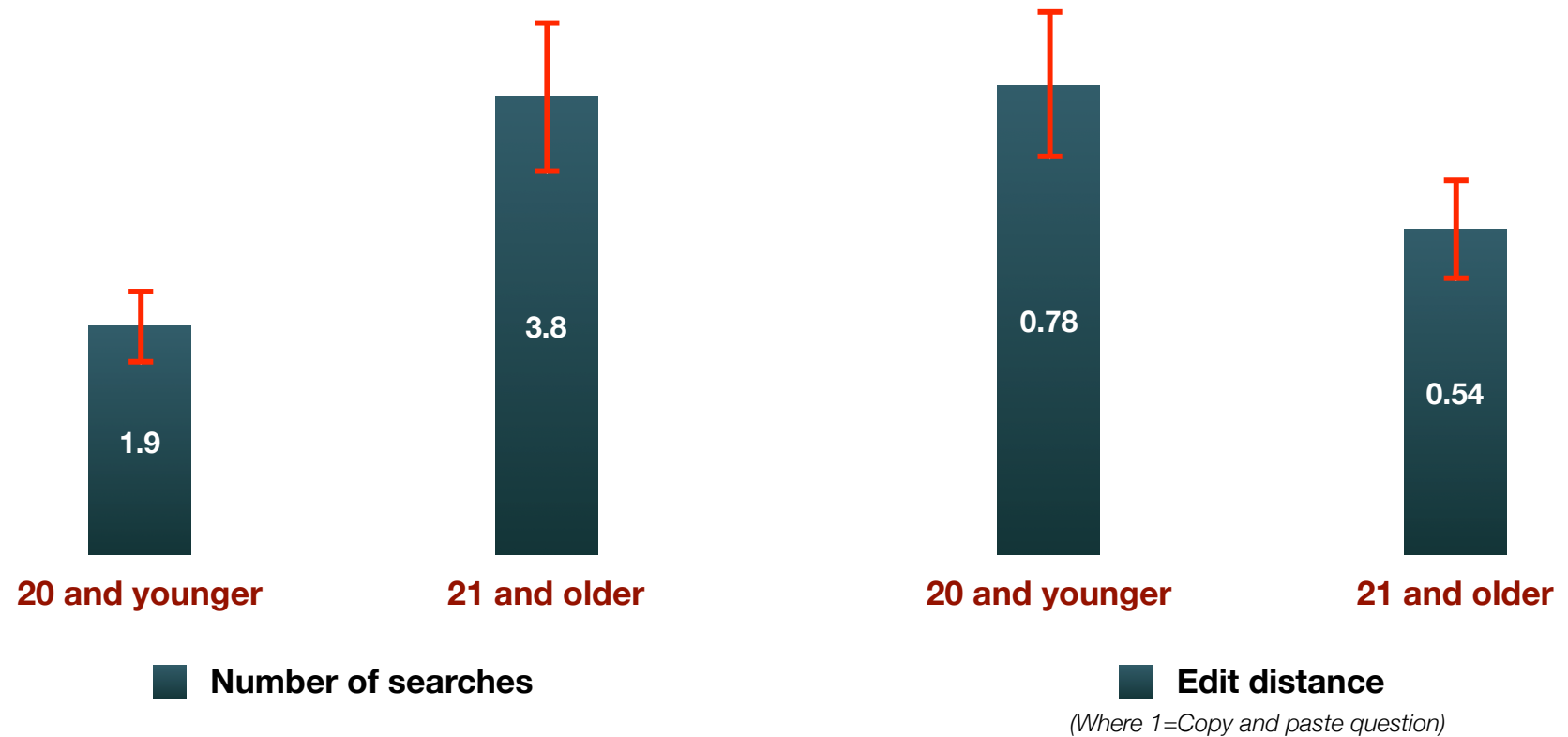
Google generation
born after 1993

Age-related differences in information-seeking

BBC Digital Revolution experiment pilot phase

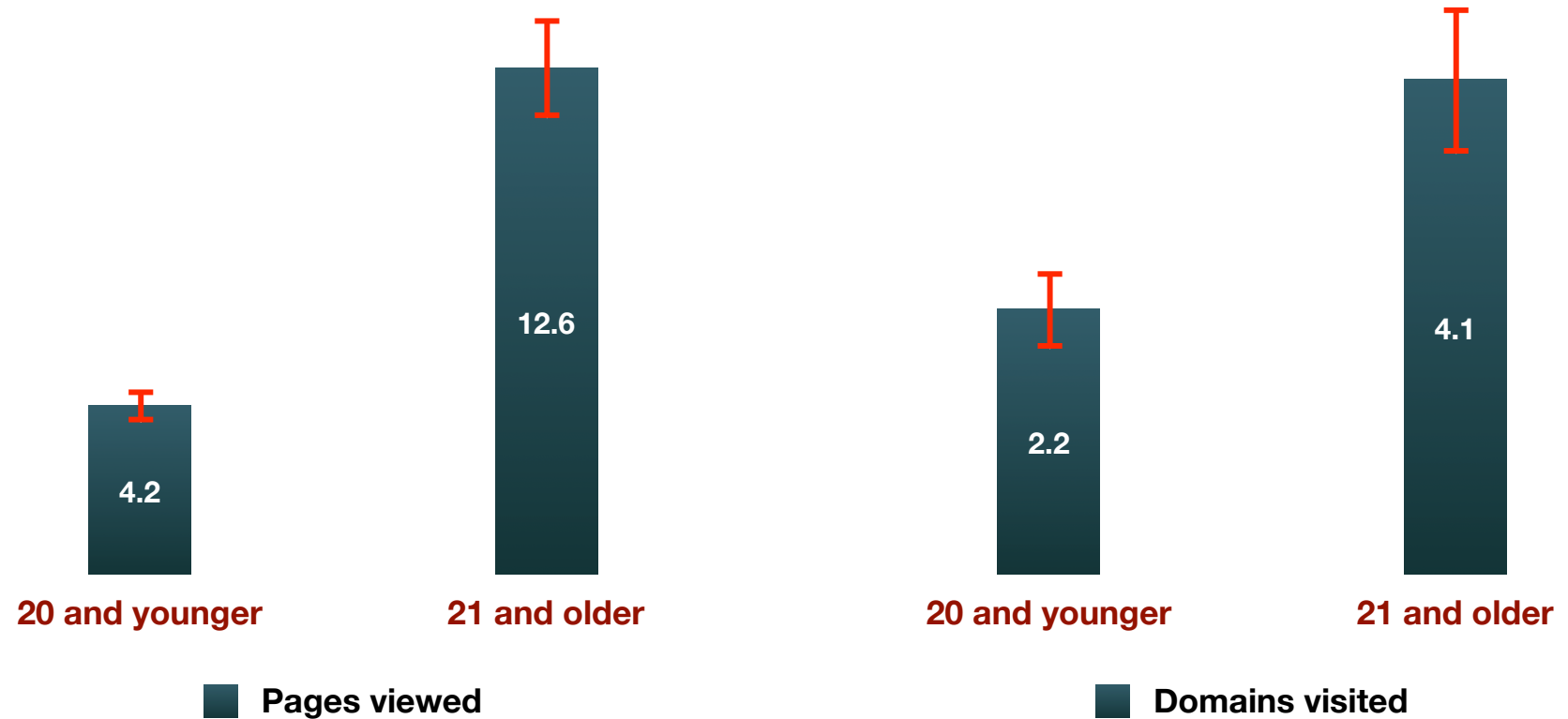
Question:

Where did the first commercial flight land?



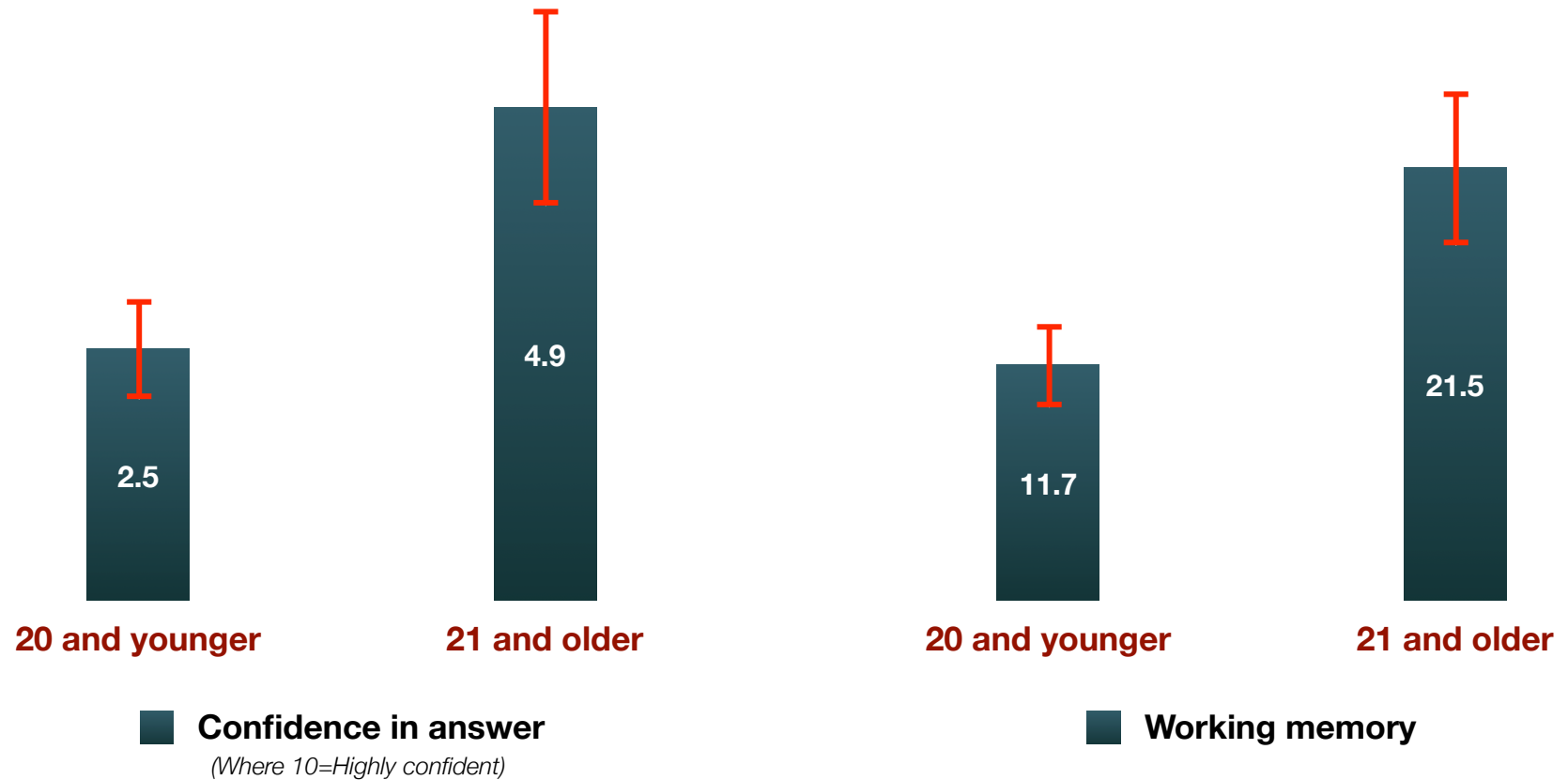
Age-related differences in information-seeking

BBC Digital Revolution experiment pilot phase



Age-related differences in information-seeking

BBC Digital Revolution experiment pilot phase



Strategic reading

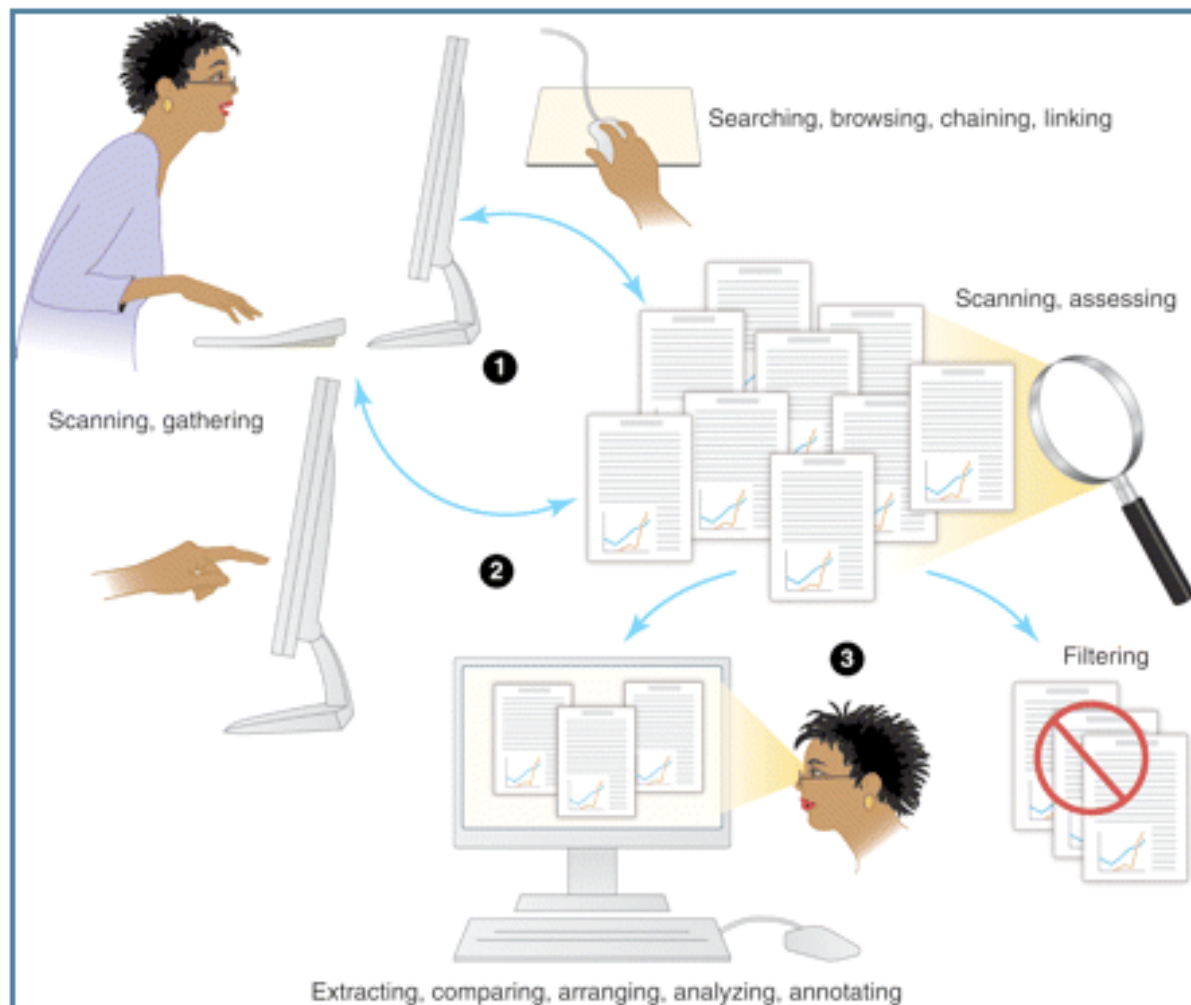
In the physical space



search
gather
browse
scan
assess
chain
filter
compare
arrange
link
annotate
analyse fragments

Strategic reading

In the digital space



search
gather
browse
scan
assess
chain
filter
compare
arrange
link
annotate
analyse fragments

Strategic reading: do scientists recognise themselves?

CIBER Virtual Scholar findings

Scientists have always strived to avoid unnecessary reading. Like all researchers, they use indexing and citations as indicators of relevance, abstracts and literature reviews as surrogates for full papers, and social networks of colleagues and postgraduate students as personal alerting services. The aim is to move rapidly through the literature to assess and exploit content with as little actual reading as possible. As indexing, recommending, and navigation has become more sophisticated in the online environment, these strategic reading practices have intensified.

Renear and Palmer, 2009: 828.

Broad levels of agreement with the above statement
CIBER survey of UK academics (n=228)

Agriculture and biological sciences	Chemistry and chemical engineering	Earth and environmental sciences	Economics and econometrics	History	Physics
80%	85%	90%	75%	55%	90%

Strategic reading: do scientists recognise themselves?

CIBER Virtual Scholar findings

The climate of intense proposal throughput and paper generation to meet targets and funding criteria will eventually push us into [this situation] driven by the lack of time for consistent reading.

Chemist, 30-39

Intensified strategic reading is a both good and bad. But it has allowed more work to be done.

Mining engineer, 40-49

In general, I agree. A scientific paper has to have an amazing and obvious draw for me to read it if it outside of my immediate field.

Agronomist, 40-49

This does reflect to some extent the changing practices over the past two decades.

Economist, 60-65

... finding papers by citations and reading only sections . abstracts of papers does speed up the [research] process

Historian, 40-49

The pressure to publish has increased the volume of literature produced in every field, with the result that it is harder and harder to keep on top of it. So, scientists use whatever techniques they can to avoid reading anything which isn't essential.

Chemist, 40-49

I rarely have time to read a paper from introduction to final conclusions. Typically I will read the methods and look at the results and then skim through the discussion.

Botanist, 30-39

The pressure to publish has increased the volume of literature produced in every field, with the result that it is harder and harder to keep on top of it. So, scientists use whatever techniques they can to avoid reading anything which isn't essential.

Zoologist, 40-49

Sounds pretty accurate. Scientists want to “do”, not read!

Metallurgist, 40-49

As soon as we stop thinking about the text as a printed page, we are freed of the primary constraint of the page ... the computer allows us to define units of text of any size and to present those units in a variety of orders, depending upon the needs and wishes of the reader. An electronic text is fluid, adjustable right up to the moment of reading. Indeed, an electronic text only exists in the act of reading - in the interaction between the reader and the textual structure.

Jay David Bolter

reading through hyperlinks rather than around bibliographic references means you are potentially reading everything at the same time, creating your own mash up

Ian Smith (UCL MA postgrad)

Selective scanning: is it *effective*?

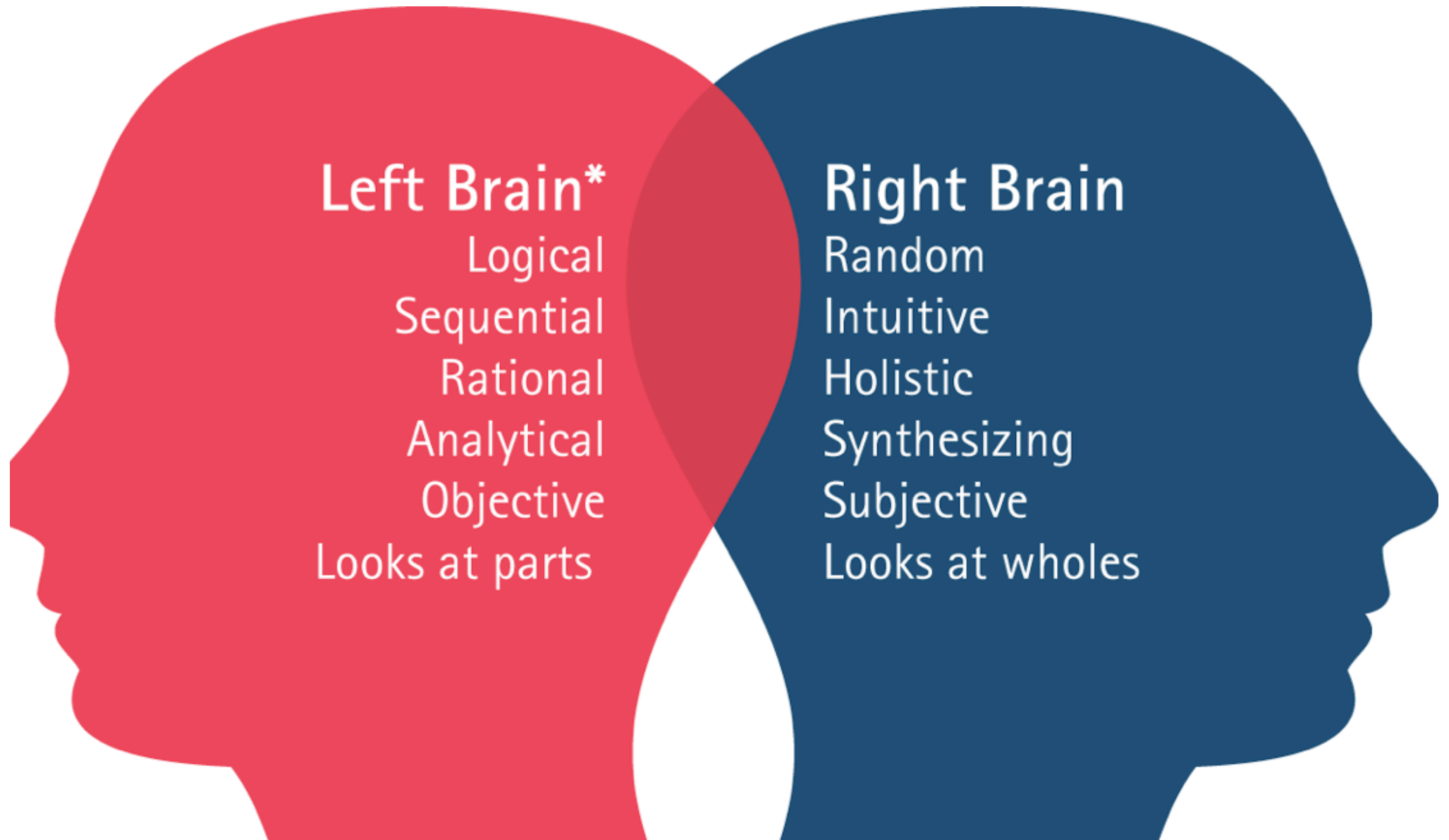
Evidence from the communications literature

Research (e.g. Eveland & Dunwoody 2002) shows that web delivery delivers less learning than traditional print. The problem is the freedom that goes with user control (“letting the patients in charge of the asylum”) in virtual spaces they do not fully understand.

While some manner of selective scanning is necessary to “tame the information tide”, the information de-selected is often the information most needed for the individual to gain a greater understanding of the content.

Users need more support!

The internet encourages over-development of the left brain



Some implications

For publishers, academics and researchers

Users need better information literacy skills (starting at school?)

Users need clearer mental maps of big complex virtual spaces so they know where they are (can we learn from supermarkets here?)

Users make their own unique journeys every time they use the web and soon forget where they are and why they are there (how do we provide context?)

Selective scanning is a fact of life, but users need more support (e.g. recommendations, tag clouds, the ability to rank output by date, citation impact or downloads?)