—————— online resources ——————

To access the resources – search on the name in italics, use the http, or search on the generic term in 'quote marks'.

Social Research Association – ethics guidelines

Singapore Statement on Research Integrity – an attempt to identify global principles

EU *RESPECT* project

Yale University – conducting research abroad – http://world-toolkit.yale.edu/research_overview

UNESCO – European cross-national research – www.unesco.org/most/ethissj.htm

International research ethics subject guide – medical research – http://bioethics.iu.edu/reference-center/ireguide

World Association of Professional Investigators – www.wapi.com

Inclusive Research Network – research by and with people with learning disabilities – www.inclusiveresearch.net

Internet research ethics – www.nyu.edu/projects/nissenbaum/ethics_elgesem.html

MORE ON THE WEBSITE

SIX

Choosing research frameworks

6.1   Studying peoples

6.2   Studying places

6.3   Studying world systems

6.4   Crowdsourcing research

6.5   Online research

Buildings have different frameworks – steel, concrete, wood – which provide the main structure. And within these frameworks different, or similar, materials are used – carpets, paint, heating. Similarly, *research frameworks*[1] (also called 'strategies',[2] 'approaches',[3] or 'designs'[4]) provide a structure within which different, or similar, data collection methods are used (C7–C13). A research design may be based on more than one framework, frameworks may overlap, and a framework can involve a combination of different data collection methods and analysis.

In the past, research design often started with a decision to use either a *quantitative*[5] (*positivist* – numbers, measurement, statistics) or *qualitative*[6] (*interpretive*[7]– words, images, meanings) framework.[8] World research can use the standard frameworks, and many methodology books explain these,[9] so they are only outlined here (Figure 6.1). But the qualitative–quantitative distinction is not now so clear-cut. New technologies are producing data that does not fit neatly into the old paradigms. Satellite images are visual data, but the images can be measured quantitatively (the number of tanks in a

war zone), or observed qualitatively (the type of tanks and their movements). More importantly, frameworks should be chosen to address the research problems in the most appropriate way, not because the researcher favours one way of doing research. There has been a belief that qualitative analysis is not viable on a world scale, except within ethnographic fieldwork,[10] but approaches have been developed,[11] and are applied to topics within international politics/relations such as gender, race, religion and secularism.[12] The internet has made online qualitative work viable and innovative.

Framework	Characteristics	Difficulties	Advantages
Action research	Cycles of interventions and evaluations to understand and change a situation.	People may not welcome interventions. A study may collapse before the cycles of planned interventions are complete.	Provides a clear focus. Applicable to training and change management.
Case studies	A study of something because it is either typical (common) or atypical (unique).	If selection of the case is opportunistic or convenience, the academic rationale may be poor.	Clear boundaries. Typical cases can be generalized.
Cross-sectional	A study of the same thing at different levels of a population or organization.	Difficult to ensure that the same things are being studied.	Multiple perspectives which increase validity.
Experimental	An intervention is tested, perhaps on one group, and comparisons are made with another similar 'control' group that did not experience the intervention.	Role play and simulation do not reflect real-world circumstances. Powerful people are unlikely to cooperate.	Useful for informing and evaluating training and pilot projects.
Longitudinal	Repeated observations of the same things over a long time period, compared with 'baseline data'. 'Cohort studies' track the same people; 'panel studies' survey a sample of a population at each stage.	High attrition rate of cohort subjects. Baseline data cannot be amended. Context changes can confound results.	Indicates effects of major policy changes. Can track development of cohorts experiencing a particular intervention.
Psychological	Assessments of the minds of individuals – personality, cognitive style, mental health.	Difficult to get data directly from actual leaders/elites, 'distance psychology' is problematic, and simulations are hard to generalize.	Insights can inform international intelligence analysis, decision-making, negotiation and conflict resolution.
Surveys (see C9)	Large-scale interviews, or postal/e-questionnaires across a population or sample.	Busy or powerful people rarely respond to postal/e-surveys or get staff to provide standard responses.	Respondents are often keen to give anonymous views about sensitive topics.

Figure 6.1   Standard research frameworks

This chapter concentrates on other frameworks which can be relevant for world research. Specialists may find the explanations very basic, but the intent is to make researchers aware of the potential of diverse ways to structure data collection, which is often lacking in international studies.[13] These data collection frameworks are explained in relation to the *conceptual framework* for this book – *peoples, places, systems*. Some studies may be *cross-sectional* looking at a topic across all three levels. *Crowdsourcing* (C6.4) and *online research* (C6.5) are also relevant across all areas. Figure 6.2 shows the frameworks that can be relevant for world research, and the website has related diagrams and templates.
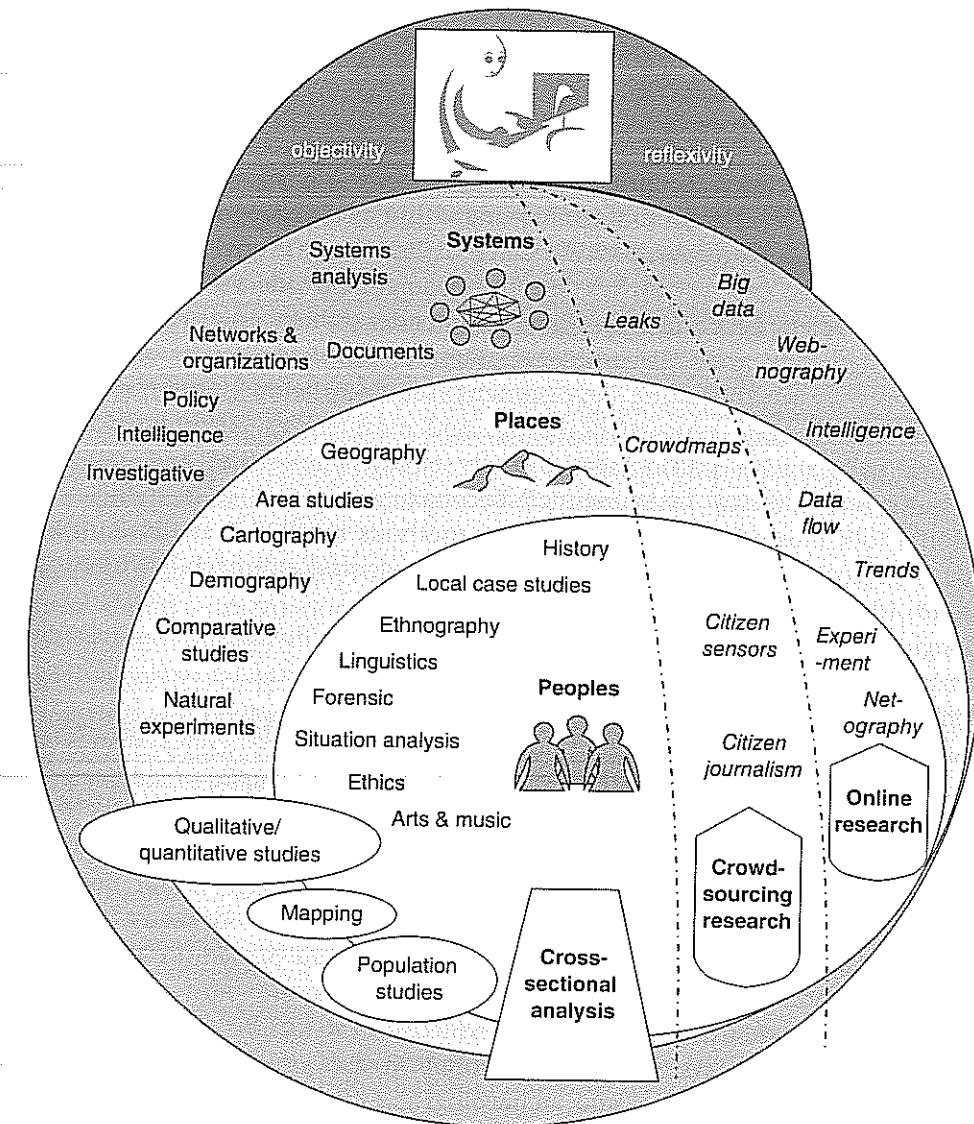

Figure 6.2   World research frameworks
http://openclipart.org/collections

6.1 Studying peoples

Local case studies are usually small-scale and focus on particular aspects of particular people's lives in a specific place.[14] They rarely concern a whole population or system, as *area studies* do (C6.2). General case study methods are relevant,[15] and the choice of the 'case' is usually because it is either *typical* (common) or *atypical* (uncommon). A case study should be structured to make it comparable with similar studies across the world, and it may be *cross-sectional* looking at the same issue at different levels in the same country.[16] The crucial factor for any case study is to define the boundaries – "What's in and what's out, and why?"

Ethnography ('culturology'[17]) is usually associated with *anthropology*, and aims to provide a holistic account of peoples or other groups, including their languages, history, place, environment, buildings, kinship, elites and objects.[18] *Fieldwork* and *participant observation* have been central,[19] and ethnographic methods can be related to critical theory[20] and reflexive thinking.[21] *Endogenous* perspectives consider things that come from within a cultural group (food taboos, house decoration); *exogenous*, concern external views (nutritional standards, symbolism). Similarly, *emic* accounts come from someone inside a group ("our gods say..."), and an *etic* account comes from an outsider ("their moral values are..."). 'Distance methods' were discussed by Margaret Mead in 1953,[22] and online ethnography is an interesting development (see C6.5 below). Forensic anthropologists[23] investigate terrorist bombings and mass graves. *Ethnomusicology*[24] (cultural/comparative musicology) initially entailed collecting and categorizing indigenous musical instruments.[25] European composers studied their own national folk music. Kodály described his methods in 1906,[26] as did Bartók in 1931.[27] Postcolonial influences led to a less Eurocentric approach,[28] which now extends to urban music and busking.

Linguistics may be part of an ethnographic study, and investigates the meaning, form and context of languages, how languages evolve, and *intercultural communication*.[29] Comparative/historical linguistics[30] can help to track the migration and interaction of peoples.[31] Language can explain power structures – India's caste system is built into language through phrases such as 'ritual pollution'.[32] *Lexicography* – using or creating dictionaries – may form part of a linguistics study. Internet and mobile phone translators add new dimensions, as do operating systems (OS) that support diverse scripts and dialects, and SMS dictionaries. But text language is not new. The *rebus* – signs or letters indicating whole words – has been used for many centuries. The *International Association for Translation and Intercultural Studies* (IATIS) provides an academic forum.

Forensic research involves collecting evidence that can be used in courts as bases for 'toxic torts',[33] cross-border litigation, ICC cases and holding governments to account.[34] Anthropologists and archaeologists examine mass graves, in places such as the former Yugoslavia, and historical photos can help. They might use private photos, and regime records, and physical evidence such as 'saw and tool marks on bone', 'surgical effects on bones (including amputations)' and 'stigmata of torture'.[35] Satellite images have been used at the ICJ, ECJ and *Inter-American Court of Human Rights* (IACHR).[36] Specific methodologies are used to organize and analyse data in a way that is appropriate for courtroom arguments,[37] and specific software is available.[38] Presenting digital evidence effectively is increasingly important.[39] Forensic evidence collection needs to be done very carefully.[40] Court records of cases can provide very reliable data for other researchers. Most importantly, researchers doing fieldwork for other reasons may suddenly find they are in a position to collect valuable evidence of abuses of power, such as children's drawings in a war zone.[41]

Situation analysis (SA) happens in many different forms.[42] It may be an independent piece of work, incorporated into planning tools such as Log Frames, or be based on *crowdsourcing* (C6.4). SA is likely to consider how people are affected by obvious factors – geographic, economic, political, cultural/social, environmental. Often *SWOT analysis*[43] (Figure 12.2), *risk analysis* (C14.3.1), and *predictive analysis* are used (C14.3). SA is often used to plan *international development* projects, or during and after *emergencies* and organizations have developed comprehensive checklists, available on the website. Project *evaluations*[44] usually compare project *outcomes* in relation to *aims*, conceptual or legal criteria, and specific *impacts*.[45] *Rapid Rural Appraisal* (RRA) is an alternative to survey-based evaluation methods.[46] *Participatory Rural Appraisal* (PRA)[47] makes greater use of tools and techniques such as *participatory mapping*, *ranking* and *prioritizing*, *community memories* and *life histories*, *timelines* and *calendars*.[48]

Ethics (moral philosophy) research compares the behaviour of people against ethical norms – international law,[49] human rights agreements, traditions, codes, professional standards.[50]

- *Theoretical* studies assess and provide the bases for ethical claims.[51]
- *Normative* studies consider the 'normal' pragmatic bases of moral behaviour.
- *Applied* studies identify how ethical standards should be used in particular circumstances.
- *Descriptive* research discovers how moral standards are actually applied.
- *Psychological* studies assess moral capacity and agency.

Global ethics considers macro-issues such as *distributive justice, environmental ethics* and the conduct of international organizations.[52] Some discussions of *global justice* assume that this can only happen within a nation state system;[53] others that the concept is older reflecting the world distribution of 'goods and bads'.[54] When analysing historical issues, distinguish between *retrospective* and *retroactive ethics* (C13.2).

Arts and music ('culture', 'media') research has been the domain of itinerant poets, musicians, storytellers and artists. Practitioners search for new materials and techniques, check the factual bases of their work, or seek inspiration. They may be *participants* through living in slums, or working rubbish dumps. The international community is interested in *cultural diplomacy*,[55] and *culture and development*, for which UNESCO has created a set of indicators.[56] Examples of *cultural imperialism* substantiate claims about colonial and neocolonial influence.[57] How governments use the arts to promote national identity can be studied by analysing policies.[58] Work within *arts management* investigates international trends and fashions, and administrative aspects including intellectual property. Journals such as *Art & Research*, and organizations like the *European Artistic Research Network* (EARN) provide ongoing discussions. A 'framework' may just be an inspirational collection of words and images, but within this a way to collect data can be devised.

6.2 Studying places

Area studies have a long tradition,[59] and usually concern major aspects of a whole country – public service systems,[60] politics,[61] religion.[62] Initially used for training international security experts,[63] the scope has now broadened.[64] Some studies are based on city data,[65] and the *Council of Europe* programme on *Intercultural Cities* provides ongoing debates. Area studies need to be structured in a way that makes comparisons with similar things in other countries feasible. The style and methodologies of area studies can be seen on sites like OECD and the *CIA World Factbook*. A series of area studies does *not* intrinsically create an international or global study (Introduction), unless they are designed to provide a world perspective. Large-scale economic and other forms of migration can make defining an 'area' population very difficult. And think about whether 'nation' is always the best unit of analysis – does it really make sense to compare China and Russia with The Maldives and St Lucia, in an index (Figure 0.2)?

Cartography – mapping and using maps – is often a basis of an area study, and many interesting methods do not require a high level of technical expertise (C11). Old maps, such as the *Atlas of the World's Commerce* (1907), can help to design a present-day study. But the most important aspect of using maps is to remember that they can deceive.[66] *Demography* relates populations to territory (C9) – the statistical changes in human (and other) populations. The *UN Population Division* (UNDESA-PD) now collects data and maintains international databases. Studies include trends such as shrinking cities,[67] *transnational migration*,[68] *mobility*[69] and *hypermobility*.[70] Ideas such as *global nomads*[71] and *Third Culture Kids* (3KC, TCK) describe the globalized identities of many young people.[72]

Comparative studies[73] are based on local and area case studies.[74] A research design is usually similar across all cases, and can be focused through a set of common questions or theory. If this is not possible, data is usually *contrasted* rather than formally compared. The design can be:

- *synchronous* – similar things happening at the same time in different places – "Immigration to EU countries in 2012".
- *asynchronous* – similar things happening at different points in history in different places – "Street children in 19th century Europe and the 20th century 'Third World' ".[75]
- *longitudinal* – cohorts of people are assessed at different points in time – "Health outcomes of children from rich and poor families in England, from 1950–2000".
- *multivariate* – statistics about numerous places can be complied into *indexes* and *indicators* (C14.4).

Numerous methods books discuss comparative methods[76] including specific advice on: politics,[77] political economy,[78] environmental politics,[79] education,[80] social policy,[81] law,[82] social justice,[83] health policy,[84] media,[85] history[86] and working with children.[87] *Descriptive* comparisons are usually based on two area studies, but rarely more than five – analysis becomes very messy across a large number of cases (C14.1).

Natural experiments entail *opportunistic observation* and *comparison* of circumstances created by 'nature' not experimenters, and can be based on *present* or *past* events, within particular places (C14.1). The *present* collapse of a government could provide an opportunity to research how civil servants cope in a crisis. Press letters, online comments to media sites such as the BBC[88] and tweets can provide a wealth of data for analysing public perceptions and the understanding of *present* events. Family histories, local official records and archaeological findings can provide evidence for *past* ('historical') studies. Historical cases can be analysed comparatively.[89] Understanding the 'collapse' of *past* societies can provide insights into the vulnerabilities of present societies.[90]

6.3 Studying world systems

Systems analysis (C12) was developed during the Vietnam War.[91] It tries to provide a holistic understanding of interconnections and interrelatedness, within a 'closed' system, such as a nation, or 'open' systems such as religions. *World systems analysis/ theory*[92] was developed by historians, including Immanuel Wallerstein,[93] in the 1970s, as it became evident that using nations and regions as the main units of analysis of world economies was inadequate. The world could be divided into *core*, *semi-periphery* and *periphery* countries.[94] *Earth systems science* investigates the interactions among the atmospheric, water, geological and biological systems. This provides the context for social science studies such as those related to the IPCC.

Network analysis[95] identifies 'nodes' of individuals or organizations who are 'tied' by interdependency arising from characteristics such as kinship, interests, beliefs, status, profession and ethnicity. Physicists are contributing to the analysis of *complex networks* in global finance.[96] Journalists use simple versions to suggest political influence,[97] and *big data* permits social media analysis (C6.5). *Supply chain analysis* may seek to make global logistics more efficient, reduce environmental impacts, or understand international crime networks (Figure 12.3). *Organizational analysis* considers the systems within organizations, by analysing factors such as context, and situations.[98] The purpose is often to gain a competitive advantage over rival organizations.[99] *Organizational network analysis* specifically studies communication,[100] and can be linked with *policy analysis* and *network analysis*.

Policy analysis[101] considers how to *make* policy, and how to *assess* policy-making, concerning public services, international politics[102] and foreign policy.[103] *Game theory* is a mathematical approach to understanding how strategic decisions might be made, when the outcome for one person/entity depends on the decisions of other people/entities in the same context.[104] Game theory does not account well for real-world dynamics such as irrational behaviour and unexpected contextual factors. Game theorists did not manage to predict the end of the Cold War.

Intelligence analysis[105] of military, political and commercial systems is used in settings such as embassies and transnational companies.[106] For outsiders, the blogs of the 'independent diplomat', Carne Ross, provide fascinating insights.[107] *Security intelligence* techniques are sometimes outlined on defence ministry websites.[108] *Centre of Gravity Analysis* determines the strengths and weaknesses of 'principal protagonists' ('the elite of the group') – friendly and hostile – in relation to critical 'capabilities', 'vulnerabilities' and 'requirements in order to act'. Human geospatial analysis can be used to assess power and populations in relation to space.[109] Software can assist. The *GeoTime* programme can map people's movements from data such as cash transactions and mobile calls.[110] *Protocol analysis* is based on the unwritten rules about the courtesies of interpersonal international relations.[111] Specialist consultancy companies provide advice,[112] and protocol manuals provide the chance to check what should have happened with what actually happened at a particular public event.[113]

Investigative research[114] is relevant to many forms of world research.[115] It usually aims to prove specific facts about state or commercial systems, and can be an element of a broader academic study. Investigative research *by* states[116] uses case-specific methods – about murder,[117] drugs dealing,[118] fraud[119] – and traditionally involves looking for *means, opportunity* and *motive*. The starting point might be a suspicion from computer-assisted profiling of a particular population.[120] Methods include electronic surveillance, informants[121] and *undercover research*.[122,123] A distinctive feature is that *groundwork* (intelligence monitoring, psychological profiling,

nurturing informants) tries to achieve well-focused investigations of the 'right people in the right place at the right time'. Investigations *of* state wrongdoing have traditionally been done by *investigative journalists*.[124] State and corporate[125] crime are significant aspects. UNESCO provides a hypothesis-based inquiry approach for journalists, which entails: starting with a few clues or facts, hypothesizing the facts that are not yet known, trying to confirm the hypothesis from open sources, questioning people who can fill in the gaps in the information found in open sources. It advises, 'Don't search for secrets. Work out what they might be from open sources.'[126] Hypotheses may come from noting what may be missing in government reports.[127]

6.4 Crowdsourcing research

Crowdsourcing (CS) can relate to all frameworks, and is probably the most exciting development in world research, because it is a way to get verifiable primary data about states, which is not controlled by states. It entails: 'an organising entity that outsources research tasks, which that entity could not achieve alone, to large groups of self-selected people (lay and expert)'.[128] This might mean managing and triangulating diverse electronic sources, such as on-the-ground mobile phone video or sensor data, satellite photos and media reports. Crowds can make handwritten records accessible by transcription and tagging. But some crowd data is hard to verify – reports of air pollution may or may not come from a particular event. Therefore the use of monitors that relay data directly to internet platforms, perhaps via smart phones, can 'take the human element out of the crowd' to transmit scientific data directly to central databases.[129]

Crowdsourced research may be organized by civil society organizations, companies or governments. Experts may contribute by systemizing and presenting complex data in easily searchable formats and analysing particularly interesting findings. Volunteers and others can be organized to contribute to data collection, data management, evidence testing (triangulation), analysis, the development of relevant technologies and dissemination. To optimize the process, research planners need to ask, what can:

- *people do that technology cannot?* (decipher unconventional handwriting)
- *people do better than technology?* (scan-reading)
- *crowds do better than individuals/experts?* (repetitive tasks)
- *individuals/experts do better than crowds?* (research design)
- *technology do better than people?* (mass data storage)
- *technology do that people cannot?* (aerial surveillance)
- *several technologies do better than one?* (robotic sample collection and spectrometry analysis)

The 'crowd' is often a social or professional network, which often parallels state structures – doctors/health, teachers/education, lawyers/justice – and so evidence about how systems operate becomes easier to track and analyse. Professionals are also likely to be seen as embodying more integrity and neutrality than activists. But crowdsourced research need not depend on the internet, and identifying non-digital methods can help to optimize the relationship between humans and e-devices. The website provides a basic CS research framework.

thinking zone: how can we crowdsource research without internet?

historical

- 1563. *Foxe's Book of Martyrs* documented 'persecutions and horrible troubles' perpetrated by the 'Roman Prelates'. This included testimony evidence from 'both men and women, whiche can and do beare wytnes', including prisoners. The 1583 edition was three times bigger than the first, and four times bigger than the Bible. *Amnesty International* now uses similar methods.
- 1858. *Oxford English Dictionary* was created by 800 volunteers who scan-read particular texts, and copied relevant quotations onto slips. Their endeavours created two tons of paper. *Wiki* texting dictionaries are now compiled similarly.
- 1890. *The Golden Bough*. Scottish anthropologist James Frazer investigated and compared instances of human sacrifice, scapegoats, the burning of humans, the sacrifice of sacred kings, and countless other violent ritual practices from around the world. He sent questionnaires to missionaries and colonial officials in many countries, including Japan, China, Africa and the Americas. In modern terms, he used social networks for a snowball survey.

recent

- 1979. *The Iranian Revolution*. When the Americans evacuated their embassy, they shredded all sensitive documents. But hundreds of students placed the shreds on the floor and numbered and indexed them manually. Local carpet weavers reassembled the pieces. The documents were published as a book – *Documents from the US Espionage Den*, which revealed the contents of telegrams, correspondence and reports from the *American State Department* and the *CIA*.
- 1985. *Bikini Atoll*. The *Hata High School Seminar* engaged in community research about the effects of *US nuclear testing*. They interviewed fishermen who had been exposed to the radiation, and mapped their stories together with family testimonies, secret government documents and evidence of how the US tried to obstruct news about what had happened.
- 2000s. *El Nadeem Centre*. Egyptian doctors working with torture victims put red dots on a map, marking police stations where torture had reportedly happened, recording the type of injuries. The same methods and 'appliances' were being used, which showed 'systematic practice' and a 'policy' dictated by senior state officials.

human–machine cooperation

Analyse examples of internet-based crowdsourcing, and consider how they might be enhanced by more use of human-based methods.

[See References for further information.[130]]

The Kenya-based *Ushahidi Project*[131] provides online CS tools which permit mass research, including election monitoring via mobile phones.[132] The *American Red Cross Digital Operations Centre*, *iRevolution* and the *Digital Humanitarian Network* organize ongoing discussion about humanitarian assistance initiatives[133] and emergencies,[134] and the *Harvard Humanitarian Initiative* is providing analysis.[135] *Tomnod Crowdrank* software helps to assess, triangulate and validate CS data that has been tagged by thousands of people. The development of $250 personal *KickSat Sprite* satellites, carrying sensors controlled from smart phones, will provide new dimensions.

6.5 Online research

Online research can also relate to all frameworks, and includes using the internet,

- as a *research tool* (email, *Skype*, search engines, internet surveys, remote viewing of real-time videos), which can increase the inclusiveness of world research (C5.3.1; Box 6.2).
- as a *source of data* (websites, blogs, social media), which is outlined here.
- for *methodological research* – online research about online research – transmission speed, nodes, specific patterns, data blocks, data management and storage.

Relevant data collection methods are discussed further in the chapters about *people* (C8) and *populations* (C9). Methods for using online data fall into two categories – researching *knowns* and *unknowns* – but there are overlaps. The website provides tools to clarify this.

Box 6.1 Big data research – terms and processes

big data – Data sets that are too big, too fast or slow, too diverse (sources and types) and/or too complex to be managed and analysed by traditional systems (climate, environmental, population movement, tweets).

distributed data – Data stored in more than one place, which needs to be combined (hazardous factories near borders + weather patterns + transport links + nearest **emergency services** + airlift capability = modelling cross-border emergency responses).

unstructured data – Data that is not clearly labelled (tagged), or systematically organized into typologies or categories (video, audio, social media postings, photos).

meta-data – data about data, often collected and stored "invisibly" (ID, time place, from smart phone use).

algorithms – a computer decision-making programme that can sort data).

text – written word or images (texts/texting – SMS messages on mobile devices).

GPS – Global Positioning System, shows 'geo-location'.

(Continued)

(Continued)

tags – labelling of digital data (topic, what/who is on a photo).

profiles – demographic (age, gender, nationality, etc.).

user ID – a unique digital identifier.

frequency – the number of times specific things happen (keywords, phone use, internet searches).

A longer list is on the website.

When searching for *knowns*, researchers know broadly what they are looking for. They know the likely data *sources*, or how to find those sources,[136] and use data that is collected specially for a particular study. They are usually trying to understand *why* specific things happen, through *in-depth* analyses of online *content*,[137] and often include researching specific *people*, or groups such as children.[138]

Searching for *unknowns* entails *big data* research – 'big' in terms of volume, velocity and variety. Researchers often do *not* know exactly where to look, or even what or who they are looking for,[139] and their studies usually concern *populations* (C9.4). The data is often *by-product* or *exhaust data*, which was collected initially for a non-research purpose (location of mobile phones, IP addresses). Programmers set up these data searches by designing algorithms – a flow chart of logical repeated actions. It is helpful for non-experts to understand the basics of these, so that research can be designed appropriately.

Big data research may occasionally provide answers to research questions, but more often identifies where to look for those answers. State *intelligence* services capture masses of digital data – emails, phone calls, texts – and then try to find something useful in relation to actual or likely events – riots, organized crime, terrorism, political dissent.[140] Researchers are often searching for *patterns* (clusters, outliers, associations),[141] *networks* (who 'shares' with who) and *trends* (frequencies, data flows), perhaps based on keywords in mass *content*. They look for indications of *what* is happening (interest in products, concern about events), and *how* large groups of people organize themselves (social networks, the leaders of e-chat), but not *why*. Understanding the basic terms and processes helps to demystify the jargon, and provides an idea of what big data research can achieve (Box 6.1).

The attraction of online methods is that they seem to link the researcher to the large and distant world easily, quickly and in a way that has never before been possible. But online methods are sometimes not well understood and over-hyped, which can lead to bad decisions.[142] The *ethical* considerations are unclear and fast-changing (C5.4), and many methodologies seem to assume that everyone in the world writes English and that the 'overwhelming majority' has access to a mobile phone.[143] Twenty percent of the people in our world still do *not* have access to electricity.

Box 6.2   Online inclusive interviewing

It is important to choose a comfortable setting to conduct an interview, so for people with visual impairment online interviews are excellent. The interviewer and interviewees are enjoying the comfort of their own homes, and sitting using their own computers which are adapted to their specific needs. And so it feels just like chatting with a friend.

هبة خليف Heba Kholeif
(International Student Advisor)

--- **main ideas** ---

Research frameworks provide *structures* to *collect* data.

The *basic frameworks* are: documentary analysis, case studies, surveys and experiments.

World research will also use *other frameworks* to collect data in relation to *peoples, places and systems, and across these:*

- *cross-sectional studies* which consider a topic at all three levels.
- *crowdsourcing research* – getting micro research tasks done by large numbers of people, often avoiding state control.
- *online research* – using the internet to search for 'knowns' to explain '*why?*', and for 'unknowns' to explain '*what?*'.

--- **key reading** ---

Buchana, D. and Bryman, A. (2009) *The Sage Handbook of Organizational Research Methods*. London: Sage.

De Smith, M.J. et al. (2006) *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools*. London: Matador.

Ford, N. (2011) *The Essential Guide to Using the Web for Research*. London: Sage.

Hammersley, M. and Atkinson, P. (2007) *Ethnography: Principles in Practice*. London: Routledge.

Hantrais, L. (2008) *International Comparative Research: Theory, Methods and Practice*. Basingstoke: Palgrave Macmillan.

Williams, C. (2013) 'Crowdsourcing research: a methodology for investigating state crime', *State Crime*, 2 (1): 30–51.

——————— online resources ———————

To access the resources – search on the name in italics, use the http, or search on the generic term in 'quote marks'.

Library of Congress – 'public domain use', country studies – http://lcweb2.loc.gov/frd/cs/

BBC Country Profiles. CIA World Factbook. OECD Country Reports. The Economist World-in-figures app – country case studies

Google Trends – shows trends in the use of keyword searches, and the level of interest for basic demographics, including country

Maps of World – www.mapsofworld.com

UNStats – international statistics

OECD Data Lab – indexes and ranking of public service systems in developed nations

Ushahidi Project – online crowdsourcing tools

American Red Cross Digital Operations Centre Digital Methods Initiative – creates and provides links to web search tools

Collaborative Online Social Media Observatory (COSMOS) – open access, integrated set of social media research tools (data harvesting, together with detection and representation of topics, sentiment, tension, frequency, information flow, geospatial location and social network analysis), with the capability to link with official curated and administrative data sources – www.cs.cf.ac.uk/cosmos

GDELT – Global Data on Events, Location and Tone – open access, 200-million geolocated global news events, 1979 to the present – http://gdeltproject.org
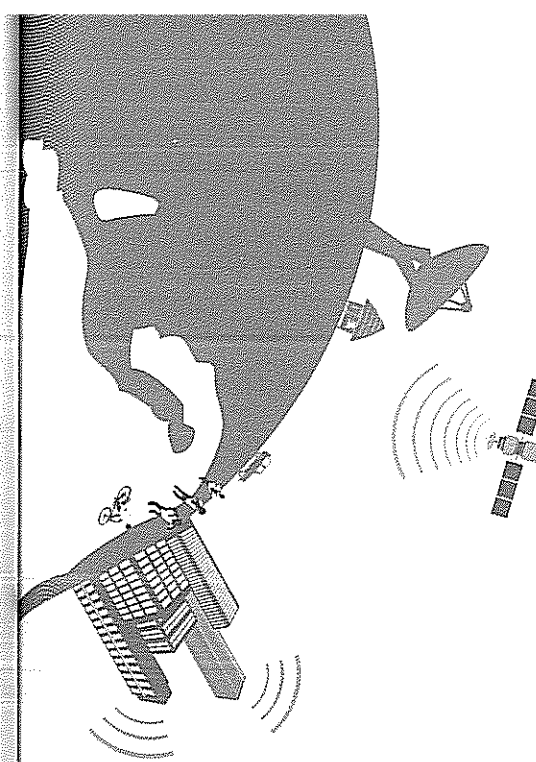
PART III

Doing
the research

Chapter 7 – The obvious problem about world data is that the world is a big place, and so innovative technologies are central to efficient *data management*. Plans for finding, selecting, collecting, testing and transferring relevant data need to be considered before starting to collect data.

Chapters 8–13 then explain the techniques for collecting data in relation to the conceptual framework of the book:

- *peoples* – identifiable individuals and groups (interviews, observation), and *population studies* (surveys, big data).
- *places* – streets, urban, rural and coastal regions (texts, objects, buildings, environment), and *mapping*.
- *systems* – administrative and organizational research, and formal documents.

Understanding how technology has developed in the past can inspire innovations, but also help with understanding and reinterpreting previous world studies.

SEVEN

Data management

Before you plan data collection, make a cup of tea. But remember that (unless you are in China) this is only possible because of an innovative data transfer technology, and theft. The invention of the *Wardian Case* – a mini greenhouse – permitted a Scot, Robert Fortune, to smuggle Chinese tea plants from Shanghai to Assam, in 1824. The result of this botanical espionage, by the *East India Company*, was the wealthy tea business in Europe.[1]

Most studies will need a plan for data management, and that starts by researching relevant and innovative methods in methods books, specialist sites and technology magazines such as *Wired*. This chapter specifically considers the *finding, selecting, collecting, testing* and *transfer* of data. But innovative IT means that these aspects are

often now carried out simultaneously. Robo-journalism systems can analyse online texts and write articles without human intervention, within three minutes of events such as earthquakes.[2]

The following chapters about data collection (C8–C13) assume a *mixed-methods* and *iterative* approach (C4.1). Not least, field trips are expensive, and when at the research site it is sensible to collect a range of relevant, or opportunistic, data in all possible ways. These chapters include doing *initial analysis* to create *findings* – the presentation of the data in a usable way with comment about possible biases and weaknesses. Chapter 14 then explains *further analysis*. But the broad aim of Part Three is to implement Marcel Proust's advice – 'to see the universe through the eyes of...others'.[3] And this does not come just from smart technologies. It needs smart researchers.

7.1 Research technologies

World research has always been influenced by new technologies to overcome problems of distance, access, scale, dispersed data, recording, and understanding others. So it is helpful to remember how technology has influenced similar research in the past, and how ideas can be reinvented and adapted to new needs. Many present-day innovations have a long history. In 2013, Google's *Project Loon* returned to using balloons to provide internet in remote regions. This uses stratospheric wind currents, much as the navigators of early sailing ships used sea currents and trade winds, and also the techniques of balloon research first used in the Boer War in 1900, which created the first *Google Earth* style maps.

A technology checklist should consider how devices might help with *access* (satellites, robots), *data capture* (life-loggers, sensors), *new interpretation* (dating, x-rays), *safety* (walkie-talkies, GPS tracking), and full *participation* of other language speakers, minority groups and people with disabilities (C5.3.1). The practicalities of using technologies abroad includes: checking the power supplies, taking spares, using robust common-brand equipment that can be repaired easily, and proactive thinking about who might repair devices in remote areas.

In sensitive settings, be aware that people might also be using devices to watch you, especially near embassies.[4] A handheld *Snoopy* device can monitor all mobile phone content, at gatherings like conferences or protests. Hotel rooms may be bugged. 'Bug detectors' are available, and there are basic strategies such as looking for small lights and listening for clicks and other noises.[5] Surveillance technologies for small lights are now cheap and available to all researchers. To be aware of what you, and others, could use, search 'spy equipment', monitor the *Surveillance Industry Index*, and check the local spyware shops.

thinking zone: how might "they" research you?

In many countries, outsider researchers may be seen as spies, sources of valuable information, or a way to make money through scams. The problem may be worse in so-called advanced counties such as America and Britain. Even friendly local colleagues may innocently report information back to line managers who are less friendly.

documents

Personal details can be used by others to send spam mail, or to create false personas, so take care with business cards, digital profiles, etc. Avoid personal identifiers that you might use on internet banking sites.

Internet

Anything you say, do or write can be compared with things about you online – university/ organization CVs, *Wikipedia* sites, publishers' sites, papers in repositories, *Twitter*, *Facebook*, blogs. Research proposals are often open access, and must not contradict other explanations about research study visits.

listening

In some countries, hotel rooms are bugged with devices in electric sockets, light fittings, and phones. Phones may be tapped, including mobile phones. At venues like conferences, all mobile phone content can be captured on small handheld devices. Metal phone boxes, or wrapping the phone in tin foil, can deter this.

talking

Anyone who you talk to may be reporting, or recording, what you say and do – even people who you know well. Be suspicious of unknown people who are very helpful or friendly.

watching

In any public place CCTV, or people, may be following you.

IT

- Any form of software or hardware may also be spyware. Be suspicious of free USB sticks, CDs, downloads, etc.
- Any document that you send/give digitally can be used or plagiarized by others. Simple defences include not putting full references on conference papers, and only putting keywords or images on *PowerPoint* slides.

online

When doing online ethnography, how do you know that those you are researching are not also researching you?

so

Brainstorm improbable, funny scenarios around these issues, in relation to your trip. How would you know if the hotel toilet were listening to you, cats were filming you, and people you are researching online were also researching you?

7.2 Data management systems

A simple data *management* system may just entail setting up files that reflect the stages of the project and headings of a final report. Many books explain how to create good systems.[6] All forms of data need accurate *labelling*, and standard templates need to be used by all members of a research team. International standards for digital data are provided by the *Inter-University Consortium for Political and Social Research* (ICPSR). Tools such as *bulkrenameutility* can rename large data sets, if necessary. For digital data, the *formats* for the final data and reporting need to be considered early, because transferring formats can be problematic, and some formats are better than others for long-term storage – PDF not *Word*, TIFF or JPEG2000 not GIF or JPG, ASCII text not *Excel*. Sites like *MullermediaConversions* can reformat data. Photo management software – *Picasa, iPhoto, DigiBookShelf* – can sort and improve masses of digital photos. If a system is set up well, writing up a study simply entails working through the files and reporting what is relevant to the narrative of the report. If the filing system is good, the report will "write itself".

Secure storage is important, because data loss can wreck a project and waste money, and ethical considerations (C5) often dictate that data must be kept safe and destroyed at certain points. Many research funders and university departments require a *data management plan* before projects start.[7] *Loss or damage* might happen due to:

- *theft or malicious damage* – How will stores and systems be locked, who will have access, and how will access be recorded?
- *impacts* from floods, rain, fire, temperature extremes – Are proofed or sealed storage and transport systems needed?
- *power surges/losses* – Are electrical problems likely, and which devices can protect equipment?
- *systems problems* – Is professional help needed to set up and manage the system?
- *human error* – Is training, and ongoing reminders, needed to minimize mistakes?
- *hacking, virus attacks, blackmail or interception* – Might someone want to steal or destroy data, and if so, how?

Responsible funders will usually be willing to pay for setting up effective data management systems. If data is politically sensitive, or could endanger anyone, keep in mind that software now exists that can overcome any encryption, and restore all 'deletes', to some extent. The most secure practice is to isolate computers (no internet or wireless), and hard drives and other storage devices when in use, and afterwards to physically destroy them by grinding or burning. Alternatively, use handwriting and manual typewriters as top secret military departments do. The *Centre for the Protection of National Infrastructure* (CPNI) provides good advice and reports about security research.

For large projects, a designated person may be responsible for data *risk analysis* (C4.3) and *curation*. This includes software updates, ensuring that data is anonymized

(removing identifiers – file names and meta-data that indicates locations and time, etc.) and removal of redundant data. Backup systems are essential, and these must be tested for compatibility, ease of use, accurate restoring, correct meta-data and network access. Remote cloud storage may be useful, through companies such as *Elephant Drive* or *Jungle Disk*, or simply by sending documents to an email file. It is usually best to separate data *storage* from data *processing* – devices can go wrong and are more likely to be stolen. Use separate hard drives, and use many small storage cards (especially in digital cameras) rather than one large card. The *Get Safe Online* project has good up-to-date advice. Security also includes thinking about paper-based data – What if a notebook is lost or stolen immediately after, or during, an interview? Can data be coded as it is being written? How will paper data be disposed of safely? Simply leaving it to soak in soapy water makes it easy to mash and often removes or smudges ink. The dried mush can then be burned.

7.3 Finding data

Locating the sources of relevant world data – "knowing where to look" – often comes from personal experience, colleagues, or literature searches (C3). For archival work the *finding aids* systemized by the *International Council on Archives* are useful. It is helpful to develop systematic ways to search for world data, because these systems may reveal data sources in surprising places. Understanding the historical context of data can provide clues, for example, that artworks may be hidden during wars.

Basic questions can structure a search for data sources, often by using big data searches (C6.5; C8.4; C9.4):

- *Who* might provide information? Most organizations have 'staff' lists, and country desk officers in international organizations such as the UN often have a responsibility to help their compatriots. Sites such as *Facebook, LinkedIn* and *Twitter* may provide basic information on named people, which can be followed up through electoral lists, registers of births, marriages and deaths, property and company registers. *Google Images* may locate people through photos.
- *When* did relevant things happen? A keyword search on *Google Trends* or *Google Books Ngram* (Box 3.1) might indicate periods of relevant events or media interest in people or organizations. Archives and specialist librarians may help with historical events.
- *Where* did relevant things happen? Searches through media sites such as *al Jazeera* or *Russia Today* may reveal regional events. Big data sites, based on media feeds, such as *GDELT* may help to locate global events.

But sources near the researcher's home may provide better information than can be found by field visits – colonial archives, museums, or immigrant and refugee community centres.

Once at a *research site*, simple *observation* – in shops, cafes, newspapers – can help locate, and understand minority and hidden groups (Figure 7.1). Local journalists and intelligence agencies address the problem of knowing where to look by ongoing *intelligence gathering* from informants, NGOs, other journalists and media, which is stored in organizational databases and archives. Companies and security services may also use *profiling* to develop sets of characteristics of those they may be interested in, which entails:[a]

- *defining problem-solution aims* – "to identify children at risk of trafficking".
- *identifying data sources* – "places attracting wealthy tourists (demand), and places with poor large families (supply)".
- *selecting out irrelevant data* – "places where children are locally available for sexual exploitation".
- *data mining* – "search for likely supply-and-demand networks".
- *verification* – "do results match on-the-ground reports (local press, NGO reports, TV documentaries)?"
- *refining profile* – "clarifying the characteristics of supply-and-demand networks".
- *decision to act* – "improve border checks at key places".

But profiling raises ethical considerations about collecting predictive data.[9] If it were possible to identify likely child-traffickers reliably, should people who fit the profile be forcibly detained or tagged, to protect children?

'Kangnam Cafe', with newspapers for the Koreans in Kazakhstan.

Figure 7.1 Finding minority groups
Photo: Author's own

7.4 Selecting data

Having found likely data sources, specific **units** of analysis then need to be selected. These will reflect the topic and focus of the study (C4.2) and also the level of the study. Studies of populations may entail researching the whole population of somewhere, but more usually a *sample* from that population (C9.3.3). This will either statistically represent a whole population (refugees, chefs) or qualitatively represent a small unique group (child refugees, immigrant chefs). But in real life, selection is often *opportunistic* – you just happen to be somewhere when something interesting occurs, and you use that opportunity to gather data (Box 7.1).

Box 7.1   Opportunistic data collection

Travel for my work as an army officer presented occasions for opportunistic data collection for my PhD. While in Iraq, Bosnia-Herzegovina and Cambodia, I was able to use personal notes, and reports and statistics from international organizations working in those countries. These provided an understanding of the context in each place. However, opportunistic data collection was potentially problematic.

- There was a risk that data sets were incomplete.
- Data that was easily available was likely to be partial and biased, particularly as it was only in English and freely shared with someone who was clearly a representative of the international community.
- Ethically, I had two distinct roles and identities – as a PhD student and an army officer – which needed to be acknowledged and kept separate depending on what I was doing.

In international research, understanding the context is particularly important. While there will be issues around any data obtained, opportunistic data collection should not be overlooked as a part of a wider approach to gathering evidence.

[Source: Johnstone, C. (2014) 'Adult education as a stabilizing response to conflict'. Unpublished PhD thesis, IoE/University of London.]

A further selection process is involved in qualitative studies, after data collection. This entails a method for extracting and organizing relevant data from notes, transcripts or observation records. The data is usually *coded*,[10] perhaps in relation to research questions, and how this is done should be explained. Software such as QDA *Miner* can help. Coding is likely to entail a *coding frame*,[11] which can be developed when the questionnaires or schedules are being devised. At a basic level, coding may just mean reading a transcript and putting the research question number next to the relevant data. At a more sophisticated level, each question might link to a research hypothesis, and even to specific "mini-hypotheses" – detailed predictions about the expected response to each question, based on literature or

theory. Alternatively, inductive or open coding permits the categories to emerge from the data, as in *grounded theory*.[12]

After expected types of data have been selected and coded, *saturated coding* can be used to assess the unexpected data that does not seem to fit anywhere. This data will either be the most important data in the study, because the researcher did not envisage it. Or it will be garbage.

Mode	Historical examples	Current examples	e-access
Recorded data was copied, noted, or curated.	Maps. Ships' logs. Travellers' notebooks. Military reports. Scientific drawings.	Fieldwork notes, diaries, photos, testimony. Video/audio recordings. UAVs (drones).	Online research. Crowdsourced data. Satellites and sensors. Screenshots.

Figure 7.2   Ways to access world data

7.5 Collecting data

There are three ways to *access* world data, explained further in C8–C13:

- Data that has already *arrived* at the researcher's "home" and is available in places like museums or archives, or through international travellers.
- Data that can be *acquired* – collected, sent, or purchased. (But consider the ethics if 'acquiring' represents stealing.)
- Data that can be *recorded* – through field notes or digital recording.

The degree of control the researcher has over data collection, on an international scale, needs thought. Obtaining first-hand *primary data* entails data collection that is controlled by the researcher, and collection often involves *fieldwork*. *Secondary data* has already been collected by someone else, collection methods cannot be controlled, but this data can be analysed in new ways, for example by museum researchers.[13] *Meta-data* is data about data, over which there is no control, such as records about computer or mobile phone use and other big data (C6.5).

Whatever the system, the aim is to create ways to find the right data, in the right form, about the right things, at the right time. Hybrid innovations that combine old and new are important. Thinking broadly about how world data has been found and collected throughout the long history of world research can inspire new ideas and the adaptation of old ideas (Figure 7.2).

Mode	Historical examples	Current examples	e-access
Arrived data just "turned up" via other people.	*Alexandria Library.* Private collections. Seals. Letters. War captives. Pilgrims. 'Human zoos'. World fairs.	Museums, archives, galleries. Documents at embassies or companies. Immigrants, international visitors, international newspapers.	TV/radio, film, music. *YouTube*. Photo galleries. Websites, blogs.
Acquired data was bought, given, taken or stolen.	Archaeological and anthropological objects. Botanical samples. Human exhibits.	Artefacts, books, merchandise. Samples from polluted sites or mass graves. Presents from friends & family.	Online auction sites. Downloaded files, photo galleries, digital archives.

7.6 Testing data

Data is just a measure or account of something, and does not necessarily represent 'fact' or 'truth', and so it needs to be tested. The starting point is to revisit the relevant aspects of ontology and epistemology (C1.1). Testing may happen as data is being collected, soon after, or during analysis.

Interview and *documentary data* might be tested by asking: who is speaking, who are they speaking to, for what purpose are they speaking, and under what circumstances?[14] Other factors include style, manner, experience and social position.[15] Certain forms of evidence may be more trustworthy than others. Data from interviewees with little public accountability, such as spies, may need greater corroboration.[16] Defector and asylum seeker *testimony* is particularly problematic, as people in these situations often need to present a specific view of why they have fled from their persecutors, or are 'countering' a prevailing view of recent history.[17] And after testimony data has been collected and analysed, respondents are rarely asked to check its accuracy.[18] Corroboration against other factual evidence is therefore vital.

Triangulation tests a piece of data by comparing it with two or more other relevant pieces of data, as a police officer might test a witness statement against other statements and mobile phone records. The aim may be to validate a significant report, but it can also be to analyse the reasons for differences. Triangulation might entail comparing evidence through different 'lenses':[19]

- *methods* – observation ('demonstrated behaviour') can be used to check documentary accounts ('stated behaviour'), interviews can check survey data.
- *respondents* –the perspectives of politicians, social workers and refugees could be compared.
- *investigators* – data collected by English and other language speakers, insider and outsider researchers.
- *theories* – post-conflict education understood through human capital, peace stabilization and human rights theories.

More usually, different types of *data* about the same things are compared – press reports, drawings, photos and political songs.

Validity describes the degree to which data and findings represent what they claim to represent. Validity is often assessed through mixed methods[20] and *triangulation*. Validity tests are often applied to interviews.[21] The *aggregation* of statistical

data ('composite data') can increase validity because the likelihood of a misleading score from a single source is greater than from taking the average score of a range of sources (C14.4). This can be tested further by a *split-half* check which randomly divides the indicators into two groups, and checks if the aggregate figure is similar for both.[22] If data is not valid, no form of testing or IT tricks can correct the problem.

Significance tests assess statistically how likely a result is due to chance (but not the degree of the causal relationship). Testing starts with a null hypothesis – that there is no difference between two variables – which the tests confirm or not. In statistical terms 'significance' implies that there is little likelihood that a result occurred by chance – usually less than 5%. But these 'significance levels' have arisen from tradition not objective statistical fact.[23] Statistical tests also do not assess the quality of data – if it was collected carelessly, that will probably not be detected or corrected by tests.

Reliability is the degree to which, if the study were repeated in exactly the same way, it would provide the same findings. If three researchers independently drew flow charts of the networks of the same transnational corporations (TNCs), would the findings be the same? Assessing reliability is often hypothetical in world research because the likelihood of repeating a study is low. But the reliability of previous studies can be assessed by, for example, identifying unclear questions or concepts, particularly with translations – does "people who voted" include people who voted but deliberately spoiled their papers as a protest? Simple *test-retest* checks can also assess reliability. In an interview a key question might be repeated throughout the interview, in different ways.

Understanding the checks used by large reputable media organizations, such as the BBC, can provide practical insights into the rapid testing of world data (Figure 7.3).[24] Professional verification is usually a mix of assessing validity and reliability, and triangulation, and a lot of common sense.

Verifying world evidence   **BBC**

Check:
- stated locations on maps. Is it plausible that someone in that place could report that event?
- images with existing images from those locations. (Is that the correct 'parliament building'?)
- images with weather data. (Was it really snowing at that time in that place?)
- audio data to see if accents and language are correct for the location. (An Egyptian would greet by saying '*Iziak*', a Jordanian with '*Kifhalik*'.)
- the original source of the upload/sequences as an indicator of date.
- weaponry, vehicles and licence plates against those known for the given country.
- previously verified material.

Figure 7.3   BBC processes for verifying social media content

7.7 Transferring data

In the 1930s, explorer Thor Heyerdahl was caught when returning to Norway from the Marquesas Islands with a gun butt carved by the artist Gauguin. Unfortunately it was still attached to the gun, which did not amuse immigration officials. But one official had the wit to ask him to detach the butt from the gun. The official kept the gun, and Heyerdahl took his precious data home. Most researchers are not so lucky.

Transferring *physical* data entails checking what can legally be exported and imported, and how this must be done safely. Many countries regulate the export of any object deemed an antique, which could be a string of beads from a market. And import restrictions can include items such as nasal sprays into Japan. Posting items may be prohibited if they are hazardous, or be very expensive and attract import duty. Some postal services, such as Britain's *Royal Mail,* will levy very exploitative charges for collecting small amounts of import duty. To minimize the seriousness of loss, post many smaller packages, each with mixed copies of data, not one large parcel. If carrying sensitive data, put it in innocent but unobtrusive places in luggage, but do not hide it in secret compartments because customs officials look for these, and obvious deceit can lead to criminal charges. A micro-SD card is probably safe in a toiletries bag, if you can find it afterwards.

Documentary data, including digital, can often be seized on the dubious grounds of national security, impropriety, or politicized regulations. This could include photos of women not wearing *hijabs* in an Islamic country. US border officials operate like Wild West bandits, and can seize anything they like, without any reason. Carrying data in seemingly innocent formats can help to avoid attention. Art may be a less conspicuous means to transfer interesting political data. A street child's drawing of police violence in Soweto during the apartheid era was unlikely to attract the attention that photographic data would, although a child's drawing of the police could also be illegal.

Postal services will provide lists of what cannot be posted without using formal *hazard* procedures, which may include batteries, human and animal remains, knives and poisons.[25] The WHO provides detailed advice about transferring hazardous materials and infectious substances,[26] and the *International Air Transport Association* (IATA) (which provides a *Dangerous Goods Manual*[27]) can help with training about classification, documentation, labelling, packing and supervision. But, of course, sometimes the most important data *is* transferred illegally. In 2013, blood and urine samples smuggled out of Syria to the British military Porton Down centre provided evidence of chemical weapons use.

Transferring data *digitally* not only applies to documents and statistical data. As devices such as 360° cameras, mini video recorders and 3D scanners and printers become more available, visual and audio data is increasingly transferrable online.

Platforms such as *Dropbox* and *Skydrive* facilitate transferring large amounts of data, and the trend towards 'open access' encourages researchers to put raw data from research sites immediately online. Toolsets such as UNICEF's *Rapid SMS* permit data collection, streamlining complex workflows and group coordination, using basic mobile phones, and the system can then present information on the internet immediately. This can assist with things like remote health diagnostics, nutrition monitoring and supply-chain tracking.

Transferring **politically sensitive** data is obviously problematic because online systems are not secure. Although seeing the 'https' lock indicates an encrypted transfer, the US, UK and other intelligence agencies can intercept and decode this. There are ways to avoid IP-address tracking.[28] The free *TOR* software encrypts data and masks the original IP address. Simpler strategies include photographing or scanning written material, and sending as JPG files, because writing on image files is harder to detect than emails or *Word* attachments. End-to-end encryption claims to be totally secure, but ultimately we never know who has set up, controls or infiltrates seemingly secure systems, and sometimes the use of secure systems attracts attention. Even governments are unable to transmit data securely, as *Wikileaks* demonstrates. Meshnet systems such as *Firechat* may be better.

New technologies are bringing significant changes to how we can transfer data around the world. Organizations such as the *Public Initiative for Information Security* support the development of more secure systems, such as *Hiding Duck*. Tracking with *Radio Frequency Identification* (RFI) means that physical data is much less likely to get lost in transit. Robots – such as the *Dolphin-cam* which can swim at 20 mph – can now stream video instantly.[29] With hydrogen fuel cell engines, UAVs, carrying cameras, will be able to stay in the air, streaming video data, for up to 30 hours. Improved 3D scanners and printers will permit the reproduction of to 3D objects quickly, from anywhere to anywhere. Conservation groups such as *Factum Arte* can digitally recreate damaged historical objects, and even small buildings, such as Tutankhamun's tomb.[30] Eventually nano-fabricators will extend this to recreating the actual materials. And if researchers still want to write and draw their field notes by hand, robot arms will be able to replicate what they do, in real time, anywhere.[31]

Data management technologies have always been significant within world research. And they have helped us to move from state-controlled companies stealing data as with the *Wardian Case*, to states and companies being deterred from stealing data with the *Hiding Duck*. Micro-cameras now come in many disguises. But always keep in mind that the people you are studying with a spying device may also be studying you with a spying device. Those blank windows at the top of US embassies and other government buildings might be hiding powerful listening and interception equipment.[32]

---

main ideas

Data management systems entail thinking about:

- *secure storage*, which protects from theft or malicious damage, and other impacts.
- *asking funders* to pay for setting up systems.
- who is responsible for *risk analysis* and *curation* of data.

Finding data entails asking:

- *who* can provide information?
- *when* did relevant events happen?
- *where* did things happen?
- is *intelligence gathering* useful?
- are *profiling* systems relevant?

Selecting data entails identifying *units of analysis*, and *samples from* relevant *populations*.

- *Probability* samples will statistically *represent* the whole population.
- *Opportunistic* and other samples are useful to study small interesting groups.
- Qualitative data will be *coded* to ensure unbiased use.

Collecting data is likely to entail thinking about data that

- has already *arrived*.
- can be *acquired*.
- can be *recorded*.

Testing data includes:

- *triangulation* by comparing different types of data, or differing views about the same data.
- *validity* tests to assess whether findings represent what they claim to represent.
- *significance* tests to assess if a result is just chance.
- *reliability* assessments to judge if similar results would be found if the study were repeated.

Transferring data is a distinct problem on a world scale:

- *Physical data* may be lost because of theft or carelessness, export or import regulations, or be expensive to post. If materials are *hazardous* they require special packing and sending in line with international regulations.
- *Digital data* is very easy to manage and send, but also very vulnerable.
- New technologies permit real-time video streaming of data, the 3D replication of objects anywhere in the world, and smart systems to deter interception.

―――――――――― key reading ――――――――――

Hesse-Biber, S.N. (ed.) (2013) *The Handbook of Emergent Technologies in Social Research*. Oxford: Oxford University Press.

van den Eynden, V. et al (2011) *Managing and Sharing Data*. Colchester: UK Data Archive.

―――――――――― online resources ――――――――――

To access the resources – search on the name in italics, use the http, or search on the generic term in 'quote marks'.

Surveillance Industry Index – tracks state and commercial surveillance systems – www.privacyinternational.org/sii/

ICPSR – international standards for labelling digital data

Trade Quotes – converting old data to current formats – www.tradequotesinc.com

Renaming files – http://renamer4mac.com

Get Safe Online – www.getsafeonline.org

Remote storage – *Elephant Drive* – www.elephantdrive.com. *Jungle Disk* – www. jungledisk.com/mote

ICPSR – international standards for file formats. Social science data – www.icpsr. umich.edu/ICPSR/access/dataprep.pdf

QDA Miner – qualitative analysis tool, including coding – http://provalisresearch.com/products/qualitative-data-analysis-soft (*QDA Miner Lite* is a free smaller version)

Public Initiative for Information Security – www.publisec.org

WHO – transporting hazardous materials and infectious substances – www.who.int/ihr/infectious_substances/en/

EIGHT

Researching people

8.1   Reading about people – documents, biographies, texts

8.2   Asking and listening – interviews, participant observation, focus groups

8.3   Watching and listening – observation, group discussions, experiments

8.4   Researching people online

8.5   Data capture

When travelling, annoying people – corrupt police, arrogant officials and inept hotel staff – are not a nuisance, they are a wonderful source of data. The starting point for most research by outsiders should be chatting to, and watching, people in airports and train stations, buses and taxis. Human data is everywhere. If you are studying child labour, your data will probably meet you from the airport bus, carry your bag to your hotel, and turn up next morning to clean your shoes.

As in Europe 300 years ago, present-day coffee shops, bars and cafes are productive research venues. Pubs might have old music boxes, or clocks, which play music recorded by pins in wooden drums, tongued metal disks, or holes.[1] Comparing performances on these early "CDs" with written scores shows that music was not always performed as it was written. For example, 300 years ago, notes *written* equally might be *played* unequally (*inégale*), like 'swing' in jazz.[2] A study of local arts could

be informed by visiting bars next to theatres, and political research could start with a cake in an underground cafe. It is useful to find out how local researchers research local people. They may have locally-appropriate methods, and realize things that an outsider would not. How else do you find out that the face iced on your cake is a South Korean political prisoner, or that the photos of the beautiful Japanese actresses are all men?

This chapter concerns research with *people* who belong to a particular place – identifiable individuals or small specific groups – *human subjects*. Data collection about large groups of people – *populations* – is explained in C9, and the study of transient, anonymous groups of people in *public places* in C10. *Reading, asking, watching, listening* and *online research* will provide a mix of primary and secondary data. *Cross-cultural* approaches to analysis are likely to be necessary,[3] but this must be done with an awareness of the dubious history of researching *others* (C2.1; C5.3). The basic methods come from *history*[4] and *ethnography*,[5] *cultural studies*, or '*culturology*'.[6] But early data about **other** peoples needs to be interpreted carefully.

thinking zone: what if...?

othering

Theories of 'othering' peoples on the basis of race and other perceived differences were created and promoted by respected elites – doctors, anthropologists, religious leaders, politicians.

counterfactuals

'Counterfactual' – contrary to the facts – analysis asks "what if..." certain things had or had not happened, to understand causal factors better.

- What if... ordinary people had *not* learned racist ideologies from elites? Would racism still exist among the masses? Would nationalistic wars have happened?
- What if... Nazi Germany had *not* put eugenics theories into practice in the death camps, leading to a worldwide condemnation of scientific racism? How might the theories of scientific racism and eugenics have developed?
- What if... Meiji Japan had not learned of the European eugenics theories? How might the ethnocentric ideologies of the Meiji rulers have developed?

futures

- What are the present, and possible future, equivalents of scientific racism?
- How can we identify 'othering', ethnocentrism and exceptionalism in our own research?
- How can the ethics of research funding try to prevent harmful research about 'others'?

Outsider researchers are inevitably shielded from certain types of local people, and so a study design should aim to *triangulate* (C7.5) a range of views from:

- *elites* – select people, often decision-makers, who can provide an *overview* of something – leaders, managers, administrators. But they may present a politicized or self-promotional perspective.
- *experts* – people who can explain *objective knowledge* gained through professional endeavours or specific skills – scientists and artists, doctors and agronomists. But their view might reflect a specific school of thinking or disciplinary outlook.
- *representatives* – people who identify themselves with a certain group and may explain *group-related experiences* – tribal or nomadic groups, young/old, male/female, professionals or manual workers. But they may only represent a partial, perhaps self-interested, viewpoint.
- *locals* – people (including visitors) who know about a particular *location* – town, village, coast – or *setting* – factories, hospitals, prisons. But they may not be able to explain their circumstances in a broader context.
- *transitionary people* – people experiencing a movement from one circumstance to another – geographical, political, personal, cultural – who have knowledge about that *change*. But during the period of change they may not fully understand what is happening.

Elites and *experts* are likely to provide *exogenous* or *etic* data; *representatives, locals* and *transitionary* people will probably provide *endogenous* or *emic* accounts (C6.1). In the 1980s, a 12-year-old South African street child might provide an insider view of prison, but would probably not be able to explain the political context of the apartheid system. Of course, the same individuals may have different personas or roles in different circumstances. Relating to these can sometimes help to win trust and gain access to people in sensitive jobs. For a gender study, going to a local football match with an army colonel may be more productive than interviewing her in her HQ.

8.1 Reading about people – documents, biographies, texts

Numerous types of *document* provide information about people, which helps to discover and compare life histories and viewpoints – press, blogs, prints and photos, letters, company or government reports. These may be *about* or *by* relevant individuals. Many countries have useful national *archives*, the homes of historical figures may have libraries, and there are countless family collections of material. Old documents can also be analysed as *objects* (C10.7). *Documentary analysis* and *content analysis* are explained further in C13.2.

Biographies provide an in-depth account of the lives of individuals, but they are often based on secondary data, and so may be unreliable.[7] *Oral histories*[8] and psychological approaches[9] can provide better empirical evidence. *Autobiography* may come from many years of methodical diary-keeping, and *diaries* themselves can be a valuable source of data.[10] *Memoirs* tend to be more personal, focusing on feelings and emotions about significant events. Autobiographies and memoirs usually disclose little that is new or interesting about the author, but may reveal a lot about their colleagues, friends and family. A selection of autobiographies might need to be studied systematically. To study Ms Z:

1. The autobiography of Ms Z is analysed in relation to a *relevant and specific time period.*
2. Friends, colleagues and others are identified, within this time period.
3. Autobiographies and biographies of those people (2) are analysed during the same time period.
4. Further data for that time period is sought through online research.
5. Analysis compares all sources, relevant to Ms Z.

Biographical analysis might compare numerous biographies, which are categorized under relevant headings (violent, non-violent), looking for factors (negative self-image, victimization, corruption), and rating the degree of these factors on a scale, to identify predictors of a characteristic (violent leadership).[11]

Images within documents need to be understood critically, especially historical ones.[12] Early 'data capture' was often set up in a studio, and the purpose was sometimes to emphasize the "primitive" nature of indigenous people, and the superiority of the elites. Looking at context can help to validate depictions. Photos showing people in the context of foreground, background and consistent natural shadows are more likely to be genuine than those showing no context, just a backdrop, or remarkably good lighting. Common sense is a good fake-detector, and with digital images, increasing the 'contrast' can help to detect changes. Unfortunately, *Photoshop* can now fix things like shadows. But *Error Level Analysis* (ELA) – on free sites such as *Imageforensics* and *fotoforensics* – can indicate digital changes and additions, and reveal meta-data that can verify the integrity of photos.[13]

The alteration of images by, or about, powerful people can sometimes be more revealing than the original image. The airbrushing of people who had gone out of favour, from formal photos, was common in Communist China and the Soviet Union. In the former Yugoslavia, Milosevic's photographers added extra people to make crowds look bigger at his political rallies – the same group appeared many times across one picture.

8.2 Asking and listening – interviews, participant observation, focus groups

Data collection through asking and listening might involve formal *interviews, participant observation* and *focus groups,* or by informally asking questions at public presentations, radio phone-ins, or during opportunistic chats. Researching people does not only mean investigating lives, but also understanding death (Box 8.1). This has many applications – in health care, counselling, forensic investigations. And as communities become more cosmopolitan, cross-cultural understanding and empathy are becoming more important.

Box 8.1   Researching death in an African context

Death must be studied within the contexts of the cultures and religions that shape people's lives.

Among healthy people, a discussion or interview about death and dying would usually be uncomfortable as it may suggest impending doom. The Igbo tribe in Nigeria prides itself as a culture that celebrates life and despises death. Likewise, in the religious context, discussing death could be seen as a curse.

However, to those who are dying, there is a more open acceptance of death. Riding on religion, the time of death and its circumstances are believed to be beyond human control, and so easily accepted, even with gladness.

Sensitivity to cultural and religious traditions is therefore important, for example asking a female to interview women.

[See: Murray, S.A. et al. (2003) 'Dying from cancer in developed and developing countries', *British Medical Journal*, 326 (368): 1–5.]

Olúgbénga Awólaràn
(International Student Advisor)

The main problem was recognized by al-Bīrūnī a thousand years ago – 'lies' (Figure 8.1). Methods books often give a lot of credence to interview data, though this may be less true in world contexts, not least because of language and cultural differences. Even if people do not 'lie', interview responses are only what someone says, about what they think about something, at a particular moment. They may have never thought about that question before, and they may have a different view tomorrow. And the presence of the interviewer can influence the responses – people may try to give "clever" or "correct" answers, and poor people may give responses that they think may benefit them in some way.

Follow-ups are therefore valuable to see if answers are consistent or if views have changed and, if so, why? This is difficult if time is limited, but it is possible to build in a follow-up mechanism, by using email, or giving stamped-addressed envelopes and cards with *Questions for next week.* A responsible local person – teacher, village head, health worker – might be asked to do follow-up interviews and forward the responses. Following up initial explanations about places and objects can also provide valuable in-depth insights about a country, which would not be evident from initial explanations.

We must distinguish different classes of reporters.

One of them tells a lie, as intending to further an interest of his own, either *by lauding* his family or nation, because he is one of them, or *by attacking* the family or nation on the opposite side, thinking that thereby he can gain his ends.

Another one tells a lie regarding a class of people whom he likes, as being under obligations to them, or whom he hates because something disagreeable has happened between them.

Another tells a lie because he is of such a base nature as to aim thereby at some profit, or because he is such a coward as to be afraid of telling the truth.

Another tells a lie because it is his nature to lie, and he cannot do otherwise, which proceeds from the essential meanness of his character and the depravity of his innermost being.

Lastly, a man may tell a lie from ignorance, blindly following what others told him.

Abū Rayhān al-Bīrunī, *India* (circa 1030)

Figure 8.1   Lies

Informal chats might include off-the-record meetings in a bar or cafe to understand the background to something. The difference between a casual chat in a bar and 'data collection' is that researchers will talk in a way that is *intentional, planned* and *systematic* – they chat with a purpose. They will have a specific *interest, research questions*, or *hypotheses*. Opportunistic chats in a shop or marketplace might not use notebooks or voice recorders, but could entail casually using the key questions of a study, for example about tourist behaviour, supply chains, or corporate corruption. But keep in mind that your interviewee may also be researching *you*. Traders may watch your face to judge your level of interest, or ask where you are staying to have your room searched or burgled while you are out.

Formal interview techniques are discussed in-depth in standard methods books,[14] including *group interviews*, and may include observing body language. Advice about interviewing in different contexts covers: international organizations,[15] political settings,[16] national contexts,[17] national leaders,[18] cross-cultural discussions,[19] spies[20] and intelligence officers.[21] Different disciplines discuss specific approaches, such as those of international relations[22] and development studies.[23] Less is written about interviews with people who have mental health problems, learning disabilities,[24] or visual or other impairments (See: International Student Advisors, Heba Kholeif. Boxes 6.2, 8.2)

Structured interviews follow a detailed interview *schedule* or *questionnaire*. This might be graduated from semi-structured to highly structured, or move from 'non-threatening' to 'threatening' questions.[25] The structure may be staged, in terms of 'the opening, the grand tour, and the follow-up', reflecting different styles – 'journalistic', 'therapeutic' (building trust and rapport), or 'investigative'.[26] In general, contentious but important topics are usually left until the end of an interview, in order to avoid the possibility that an interviewee may walk out before answering most of the questions. The format is often complicated if it entails *routing* – "If no, then go to..." Tablet or laptop-based interview schedules can make this easier, and when answers are in-putted initial analysis may be immediate.

Semi-structured interviews permit flexibility yet maintain focus, but it is hard to maintain consistency if they are done by more than one interviewer. These might explore *life histories*, or memories of a significant event, in-depth. Alternatively, mixed approaches might be used by 'tandem' interviewers who switch roles strategically – one notes body language while the other asks questions.[27] How will interviews be *recorded* – notes, audio recorders, videos? There may be no second chances, and so backup recorders may be vital. If the setting has a lot of background noise, or the interviewee does not speak the interview language clearly, lapel microphones can help.

Investigative interviews often use psychological methods, and in forensic settings interviewers need a good knowledge of process and relevant jargon.[28] *Cognitive interviews*[29] use techniques to enhance memory and increase recall by people who appear unconfident when replying to questions. Techniques include reconstructing context, insisting on complete unedited descriptions, recalling events in forward and backward timescales, changing the perspective of the recall – seeing from a different angle or as another person. A central principle is that each interview needs to be planned in a way that is compatible with each interviewee, especially traumatized people such as defectors.[30]

Experienced interviewers may use techniques for *extending interviews*. Having asked a few key questions they might:

- *comment on an object in the interviewee's room* – a picture, vase, carpet – which may lead to a discussion about where it came from, which provides data about social and family networks.
- *build on what the interviewee has just been doing* – if they have just come from a meeting or field visit, ask about it.
- *ask for advice* – "If I (my son/daughter/friend/students) wanted to get a job like yours, what would you suggest?" – "What would you put in a curriculum for management training?"
- *use flattery* – "Your staff all seem so cheerful. Why is that?" – "I don't think there is anyone else who could explain this properly." – "I think you understand this better than anyone else."
- *mention mutual, or impressive, friends or colleagues* – "I also studied English at that college, did you know...?"

Interviews may be optimized if the interviewee believes she or he will be compared with others. Other interviews can be mentioned to prompt or provoke more discussion – "I believe that the *Imam* does not completely agree with you..." It is always worth asking, "Who else would you suggest I talk to?" because this reveals networks and provides the basis for negotiating access. Providing feedback of initial analysis may be an effective means to gain more data and improve accuracy, because at that point interviewees have invested time in the work and have an interest in ensuring its accuracy.

Focus groups[31] provide a compromise between interviewing and a full survey (C9.3). They can be done online,[32] utilize a range of technologies,[33] and be used internationally for topics such as health[34] and media studies.[35] A facilitator will ask basic questions, and perhaps use prompts to elicit further views.[36] Group dynamics can be noted – how members influence one another, and how discussions evolve. For some groups, a more visual *non-verbal* method might be more appropriate. Reliable recording is essential, as it is not possible to repeat anything. A simple focus group might be recorded with two voice recorders. A lone researcher might use a *smart pen* which links an audio recording to written notes. A captioner (court recorder) could use a stenotype machine linked to a laptop. If clear and comprehensive records are needed, individual radio microphones can be linked to different channels on the recorder. Different video cameras might film specific aspects such as group-shots, leadership, body language, males and females. Another researcher might note non-verbal body language, and things of especial interest in a notebook (Figure 8.2).
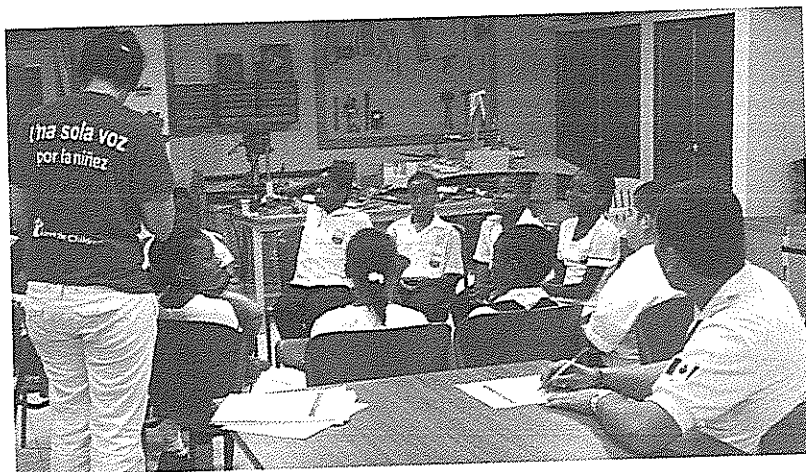

Figure 8.2 A focus group

Photo: Angela Latorre (with permission from participants)

Translation and interpretation are often main considerations. Phone-based translation services and smart phones are useful for the basics, and this includes *Video Remote Interpreting* (VRI) using sign language for people who have hearing, speech or other language impairments.[37] Bilingual research assistants are likely to do a better job if they have a good understanding of the aims of the whole project. For distance interviews, sending interview questionnaires to interviewees who speak other languages might help them to prepare clear responses. Or it could lead to sanitized answers prepared by lawyers. A compromise is to outline the areas of questioning – "What you learned from your sitar teacher in Bangladesh?" – but not the exact questions. Online and smart phone translation software creates new opportunities.

Transcription of recorded data is expensive. For a linguistics study, full transcription may be essential to understand the detail. But when general interview data contains a lot of irrelevant material, a researcher might just note the counter numbers on the recorder when something relevant is said, and then only transcribe those parts. If translators or interpreters are used, how will they be briefed, and will their transcripts be double checked by back-translation? This was understood 1000 years ago (Figure 8.3).[38] *Voice recognition software* (*Word* – 'accessibility', *Dragon Naturally Speaking*) can create transcripts, but it needs to be trained to recognize specific voices.

> ...we have sometimes written down a word from the mouth of Hindus, taking the greatest pains to fix its pronunciation, and then afterwards when we repeated it to them, they had great difficulty in recognising it.
>
> Abū Rayhān al-Bīrūnī, *India* (circa 1030)

Figure 8.3 Back-translation

Interviews with *elites* and *experts* usually entail a brief one-off chance to meet a busy person at an inconvenient moment, and present distinct difficulties.[39] Not least, people who have power on a global scale may try to control an interview.[40] The strategies for gaining *access*, *"surgical" questioning* and *prolonging interviews* are explained in *Researching Power, Elites and Leadership*.[41] Interviews with politicians may amount to extracting specific 'memoirs' rather than information.[42] Promising feedback of initial analysis to managers may be an effective means to gain more data and improve accuracy, because at that point they have invested time in the work and have an interest in ensuring its accuracy. Preparation is vital. Asking busy people obvious questions that are answered on their website is not a good way to start an interview.

Talking and listening with *representatives, locals* and *transitionary* people may not be so pressured, but the techniques may need to be more sophisticated to ensure inclusion of vulnerable or minority groups, children[43] or women. Indirect methods may work best with vulnerable people, who may have had bad experiences from answering formal questions. Thematic drawing (draw and tell) is an effective methodology, and people "tell a story about the drawing" rather than answering direct questions.[44] Asking children to do thematic role play with glove puppets – made quickly from plastic bags, tape and paper cups – can also overcome inhibitions (Figure 8.4).[45]

Figure 8.4 Using puppets
Source: Williams, C. (1990) 'Street children and education' – http://etheses.bham.ac.uk/698

Certain people, cultural groups, contexts and places may require greater *sensitivity* (Box 8.2), and this is not only because people are vulnerable. In his research among Taliban leaders, Matt Waldman used a 'divide and rule' strategy – 'all interviewees were contacted and interviewed separately...none is based in the same district as another, and none disclosed to comrades that they were being interviewed'.[46]

Box 8.2   Sensitive interviewing

In Egypt, talking about anything related to sexual experiences is considered a taboo. This is even more difficult when the interviewees are visually impaired, as society has very low expectations for these girls as wives-to-be. Educating them about sex is very rare, and families, especially mothers, do not discuss this.

So I avoided asking the interviewee directly about herself. I used the third person, or I tried to ask as if I were asking their general point of view about sex education, or if it is different for a female with visual impairment, or how society views her as a wife. I asked a mother as if I were talking about her view of *other* parents of girls with visual impairment.

And things changed after the 25th of January Revolution, because before that Egyptians lived in an authoritarian regime, and discussing politics was not allowed. Now, everyone is discussing politics. But because there is a strong attack on the Muslim Brotherhood from the media, people who are pro the Brotherhood are afraid to declare this as they are scared of the criticism they will face from the general public.

Heba Kholeif هبة خليف
(International Student Advisor)

Voice Over Internet (VOI) and *phone interviews* require careful preparation,[47] because misunderstandings are hard to remedy at a distance. Not least, check

time zones. Like any interview, the opening 'warm-up', to establish rapport, needs to be planned. If *Skype* or other video systems are used, the question of what to wear may be important, because the image will probably be small and poor quality. Similarly, check what is behind you – a glamour poster may not create the right impression. If possible, use a separate camera and microphone, because the quality is better. Microsoft has developed fully interactive translation software that can be used on VOI systems. Photos and *PowerPoint* slides can be used with VOI, as a basis for questions. An online interview can have gaps – for minutes, hours, or weeks – while interviewees find helpful materials or check facts. Similarly, it is possible to do experiments online, and to ask interviewees to talk to colleagues, family or others.[48] Short *email questionnaires* can work well. But, unless it is certain that the respondent will provide a long reply, put the questions in the email, visible in one email screen. Attachments or a seemingly endless list will probably not get a reply. Online surveys are outlined in C9.3.

Life-logging – automated diary-keeping – provides an alternative to interviews. Devices such as *Autographer* can capture visual and audio data at intervals throughout a day, automatically about people with interesting lives – traders, refuse collectors, farmers. Alternatively, they might become research assistants – 'human sensors' – and record specific data occurring during their daily life – tourists, fishing, police conduct. *GPS tracking* devices can be used, but this requires full explanation and consent from those who wear them. *Walkie-talkies* can permit assistants to send commentaries directly to researchers, and to keep in touch for safety. 'Electric tattoos', in the form of tiny microphones stuck to the throat, may soon be able to transmit speech wirelessly, and eliminate background noise. This creates the possibility of very unobtrusive real-time data transmission. These techniques obviously require specific training, ethical discussion and risk analysis (C5.4).

Advances in *robotics* and *natural language processing* are presenting new possibilities.[49] Robots are able to have interactive discussions – to teach languages, provide therapy,[50] chat with astronauts,[51] monitor elderly people. Robot-mediated interviews have been tested with children, using a familiar 'schedule': establishing a rapport, asking for free narrative recall, asking questions, closure and thanks.[52] Independent robot interviewers present many potential advantages. Robots will keep to their script consistently across countless interviews, and work in many languages. They could route yes/no answers quickly, use Likert and other rating scales, record video and audio, and use voice recognition to transcribe and translate answers. The data could be sent wirelessly to a cloud platform, in real time. Robot interviewers can smile, wink, nod encouragingly and sing songs. They have infinite patience and stay polite, appear confidential, and require no training after being programmed. Robots are much cheaper than humans and, if they disagree with you, then you can just switch them off (Figure 8.5).

Figure 8.5 A robot Interviewer (Japan, 2013)
Photo: Author's own

8.3 Watching and listening – observation, group discussions, experiments

Observation[53] – *watching* how people behave and *listening* to what they say and to the music and sounds they make[54] – provides the chance to analyse language[55] and body language.[56] Observation can also happen during interviews (C8.2), group discussions and experiments,[57] and might include reflection on how the experience affected the researcher. As with interviews, *translation* and *recording data* need to be considered. *Observation schedules* provide a framework for logging data, and use time-frames and shorthand symbols. They might note the frequency of certain behaviour in one-minute blocks, for example how often police take bribes from motorists at checkpoints. 'Rapid' methods are evolving.[58] Observations are usually recorded as *field notes*, which are often in digital form.[59] But in sensitive contexts, some form of *encryption* of notes and sketches may be necessary, in case they are seen by participants. This can use very simple signs such as circles, lines, crosses for different types of people or activities.

Within *ethnomusicology* (C6.1) listening and observing are complemented by *listener research*. Recording devices not only capture data, but can present audio

material to elicit responses. A study might play respondents folk or pop music from other countries, to discover how they perceive it. Researchers from countries that were formerly the "object" of study are now researching how their own cultures have international influence, for example the influence of Kazakh music in Europe. Mobile phone and other audio recorders make basic ethnomusicology feasible as part of any fieldwork.

Participant observation[60] has been central to ethnography,[61] and is based on intensive, usually long-term, involvement, which may range from partial to total. This raises ethical and risk concerns (C5), which entail trade-offs between deceit, personal safety and over-influencing the behaviour of those being observed. A range of standard methods are used, and data collection may be overt, covert, or 'unobtrusive'.[62] *Visual research* may consider accessories such as clothes, hairstyles and electronics.[63] For an outsider, blending in with a group is not always easy, and there are sometimes no simple solutions, except to be aware of how others will see you.

New technologies are creating many opportunities for *remote observation*, and the principles of traditional forms of observation can still be applied. Parliaments, war crimes tribunals, shareholders meetings and many similar events are webcast in real time, and have archives, podcasts and transcripts of proceedings. These provide the chance to observe people in challenging situations in great detail. Observation can include question–response styles, the use of prepared versus spontaneous responses, media skills and obfuscation. *Emotions analytics* software can use 10–15 seconds of speech, irrespective of language, to analyse *how* things are said, rather than *what* is said, to assess mood, attitude and emotional personality.[64]

8.4 Researching people online

Online research (C6.5) can use the internet as a research *tool* (examples are in the relevant sections above), or as an *object* of research (discussed here and C9.4).[65] Online studies about people are usually small-scale, feasible for single researchers, and are explained under headings like 'digital anthropology',[66] 'virtual ethnography',[67] 'online ethnography,[68] 'netography',[69] 'multi-sited ethnography'[70] and 'social media research'.[71] Studies might focus on behaviour that is uniquely online, such as online grooming.[72] Data collection may simply entail capturing relevant material for later analysis using screen shots, or researchers may make notes like traditional fieldwork notes. The internet also creates the opportunity for large-scale cross-national/cultural web-based *experiments*, which are mainly in the field of psychology.[73] Simple online tools are available, and the *Web Experiment List* helps with recruitment and lists experiments. Open source software is increasingly available, in specific fields

such as *international development*.[74] Examples of online questionnaires are on the website.

Relevant *online texts* may be found under headings such as 'staff profiles', 'management structure', or 'chair's report'. *Wikileaks* proposes a simple methodology for accessing and *crowdsourcing* analysis (C6.4) of its online data:[75]

- search for events you remember that happened, for example in your country.
- browse by date or search for an origin near you.
- pick out interesting events and tell others about them.
- use *Twitter, Reddit,* mail – whatever suits your audience best.

Online texts are not always what they seem and the sources need to be verified through checking the web address and contact details against other information. Comparative research can be based on keywords and likely names (Figure 8.6).[76]

As I got older, that gut instinct – that America is the greatest country on earth – would survive my growing awareness of our nation's imperfections (Missouri, June 2008).

I believe in American exceptionalism, just as I suspect that the Brits believe in British exceptionalism and the Greeks believe in Greek exceptionalism (NATO, Strasbourg, April 2009).

I believe we should act. That's what makes America different. That's what makes us exceptional...let us never lose sight of that essential truth. (Speech to the Nation, 10 September 2013).

Some may disagree, but I believe America is exceptional, in part because we have shown a willingness, to the sacrifice of blood and treasure to stand up, not only for our own interests, but for the interests of all. (Speech to UN General Assembly, 4 September 2013).

Barak Obama, President of the USA

It is extremely dangerous to encourage people to see themselves as exceptional, whatever the motivation...We are all different, but when we ask for the Lord's blessings, we must not forget that God created us equal. (*New York Times*, 20 September 2013)

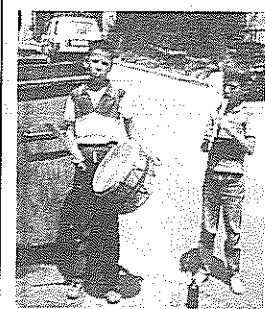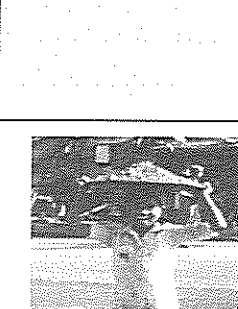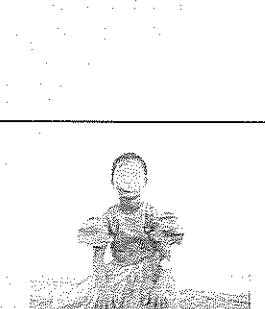Vladimir Putin, President of Russia

Figure 8.6   Comparative exceptionalism

Simple searches of *big data* can sometimes provide clues for further in-depth research, and can be very easy to do. A name search on *Google Trends* or *Ngram* (Box 3.1) may indicate periods of online activity or media interest. Looking at the metadata for simple searches should show the number of internet hits for a particular name and phrase. Sometimes, studying the *human networks* around online networks may be more significant for addressing particular research questions, but these may be hard to access and understand. Some user groups, such as hate-groups, may use chat rooms as a perverse game, and the content may reveal little about what they actually think or do. How could the disability comments around the Korean ILBE site be interpreted and researched? ILBE purports to be a conservative political group, and users call themselves 'retards'.[77]

8.5 Data capture

The technology for *data capture* about people is increasingly simple, but the ethical, legal and safety considerations are increasingly problematic (C5.4). Photographic and other electronic data permits retrospective *categorization* and *typologies*, which may not be evident just from observation, but each piece of data raises unique ethical considerations (Box 8.3).

Box 8.3   Photographing street-working children

Photographs of children working provide data about dress, health (note legs), hygiene (note clothes), occupation, behaviour, relationships and networks, abuse, and micro-economics. But should working children be paid for their time? Is it different if they also provide a "service" such as entertainment?

(Bombay/Mumbai, street circus. Istanbul, Roma musicians)

Working boys are conspicuous; girls are not, especially sex-workers. The boy said he was working on construction sites with his family. The girl said she wore a "party dress" during the day, and changed into jeans to go home in the evening.

If the boy were paid for an interview, because it was done during his "working time", should the girl be paid for the same reason?

(Bombay/Mumbai)

Typological analysis: Photos help to place children within analytical typologies, such as:

- totally abandoned, *of* the streets, *on* the streets, in families.

Note: potentially violent adults may be watching and using street-working children, and can pose a safety hazard for researchers and children.

Procuring appropriate devices is not simple, and comparisons can be made using checklist headings like: *device, capabilities, examples, advantages, problems.* Relevant equipment is listed and explained under these headings on the website. But keep in mind that the use of any devices could be used by police as evidence of spying or terrorist activities. Although small hidden devices are less likely to attract attention, if found, they will appear more like spy devices. Check country law about the use of all devices. Remember, the people you are watching may also be watching you.

main ideas

When researching identifiable **people**, consider methods from *sociology, anthropology/ ethnography, elite studies, biography* and *investigative* journalism. A triangulated sample of people might reflect:

- *elites* – select people, decision-makers, who can provide an *overview*.
- *experts* – relevant professionals and skilled people who can provide *objective knowledge*.
- *representatives* – people with group identities who can provide *insider* perspectives and explain things about their group.
- *locals* – those who can provide *first-hand information* about a familiar location or type of place.
- *transitionary people* – people moving from one state to another, who can *describe that change*.

Data collection methods include:

- *reading* – biographies, websites, blogs, CVs, 'staff profiles'.
- *asking and listening* – using informal chats, formal structured or semi-structured interviews, with individuals or groups, which may require interpreting from other languages, including signing.
- *watching and listening* – observation of speech and behaviour, and the music and sounds people make, to study language, body language and culture.
- *online research* – ethnographic examination of people, experiments, big data searches, using the internet.

key reading

Barbour, R. (2007) *Doing Focus Groups.* Thousand Oaks, CA: Sage.

Gillham, B. (2008) *Observation Techniques: Structured and Unstructured Approaches.* London: Continuum.

Jorgensen, D.L. (1989) *Participant Observation: A Methodology for Human Studies.* London: Sage.

Makagon, D. and Neumean, M. (2009) *Recording Culture: Audio Documentary and the Ethnographic Experience.* Thousand Oaks, CA: Sage.

Weiss, R.S. (1995) *Learning from Strangers: The Art and Method of Qualitative Interview Studies.* New York: The Free Press.

online resources

To access the resources – search on the name in italics, use the http, or search on the generic term in 'quote marks'.

Emotions analytics – assessing mood and personality from speech intonation – www. beyondverbal.com

Dragon Naturally Speaking – voice transcription software – www.nuance.co.uk/dragon/index.htm

Digital photo analysis – Imageforensic.com – fotoforensics.com – www.imageforensic.org/

Qualitative Researcher – recording techniques – www.qualitative-researcher.com

Evaer – recording VOI (*Skype*) data – www.evaer.com

Recording radio – http://radio.about.com/od/recordstreamingaudio/

MORE ON THE WEBSITE