MSc Research Skills

Topic: Ethics & professionalism in science

D G Rossiter

University of Twente.
Faculty of Geo-information Science & Earth Observation (ITC)

May 2011

Topics

1. What are ‘scientific ethics’?
2. Fraud
3. Authorship
4. Intellectual property and fair use
5. Professionalism
6. The social responsibility of the scientist

Topic 1: What are 'scientific ethics'?

Ethics: correct behaviour within some social setting.

Scientific ethics: narrow and wide senses:

1. narrow: scientific **procedures**: **rules of conduct**

2. wide: scientific **activity in general**: **relation to society**

Scientific ethics (narrow sense)

Two main principles:

- **Honesty**: there is a 'true' real world, only by honest science can it be successfully understood.
- **Credit** for work performed: **reputation** is the **currency** of the scientific world

Basis of scientific ethics (narrow sense)

- ethical behaviour should **advance** the overall scientific project
  - advancement of knowledge of the world
- **internal** to science, not from ethical systems
- advancement of the **human** "project"??
  - If science advances, so does humanity??

Basis of scientific ethics (wide sense)

- **External** to science
- Societal values (humanistic, religious, ideological): what is ‘good’?

Topic 2: Fraud

Scientific **fraud**: any action which wilfully mis-represents the truth

Note **intent to deceive**; sloppy work is poor science but not fraud.

Three types, in order of seriousness:

1. **Fabrication**: making up data, lying about procedures;
2. **Falsification**: manipulating data to obtain a desired outcome;
3. **Plagiarism**: taking credit for someone else's work.

Data fabrication

Fabricating data: inventing data or lying about the procedures by which it was obtained.

This is the **cardinal sin** against science, because it can **never be un-done**.

Examples:

- filling in survey sheets without actually making field visits
- over-interpreting a survey response or field observation
- inventing lab. results without actually doing the procedures

Without accurate primary data, the entire research is invalid.

Mistaken facts vs. mistaken ideas

> "**False facts** are highly injurious to the progress of science, for they often endure long; but **false views**, if **supported by some evidence**, do little harm, for everyone takes a salutary pleasure in proving their falseness and when this is done, one path towards error is closed and the road to truth is often at the same time opened."
>
> – Charles Darwin, 'The Descent of Man' (1871)

This could be better written "false **so-called 'facts'**" . . .

Data falsification

Falsifying data is **manipulating** actual data to obtain a desired outcome.

Several forms:

- **omitting** 'inconvenient' observations (e.g. reporting only successful experiments)
- silently **changing** data values to more 'reasonable' ones

Discarding data

Under certain circumstances, data may be discarded; but:

- always **explicitly mentioned**, at least in the lab. or field notebook, probably in the methods report;
- based on **objective criteria** that are **equally-applied** and **justified**

Note: if data is discarded, the work may now refer to a smaller **population** than planned (e.g. agricultural soils vs. all soils in a region)

Discarding data during sampling

Example 1: a planned soil fertility sample was found to be located in the middle of an irrigation ditch; this can be discarded because it's not representative of the population being sampled (i.e. agricultural soils).

This is on the basis of criteria defined **prior** to beginning the sampling.

Example 2: a respondent in a household survey seems clearly to be mentally ill and delusionary. Record his or her answers, but add a note about their mental state as you interpret it, and then state that this response was discarded for the reason that, in your opinion, the respondent was not reliable. Another researcher can still make use of your primary observation if they disagree with your assessment of the respondent's state.

Discarding data during analysis

'Outliers': data points that don't fit an overall pattern. . .

. . . but these may be the **most interesting** and give the **most insight**

- Must be reported in the raw data
- Criteria for eliminating must be clear and consistent
- must argue that they are not part of the **population** being analysed.
  - Poor technique (but how do you that know only this sample was affected?)
  - Poor record-keeping (reflects poorly on your technique, but at least you are admitting it);
  - From a markedly-different site that is not included in the population you are studying.

An obvious **recording error** (e.g. missing decimal point) may be corrected with no further observation.

Manipulating raw data – why?

It may be necessary to **adjust** raw data to **correct** for inconsistencies, e.g.,

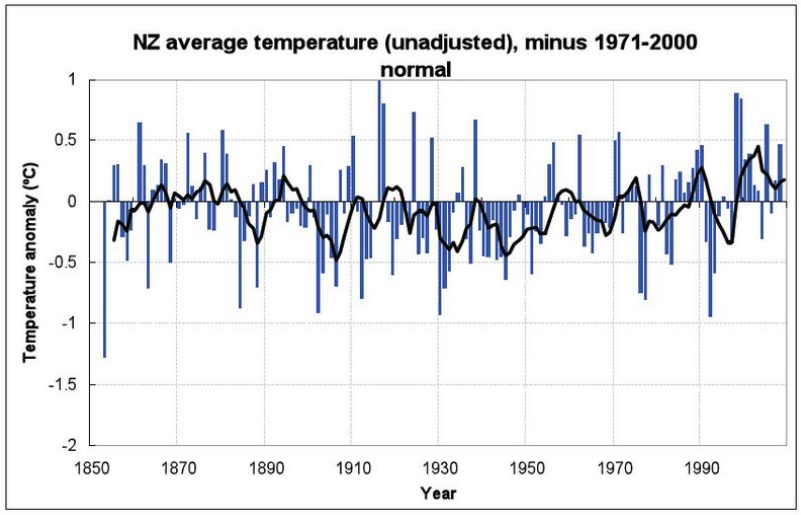
- different instrumentation or analytical methods to measure the same thing
  - within the same experiment
  - change in procedures over time (time-series, e.g., weather or soil monitoring)
- different operators (researchers) measuring the same thing
- the aim is to achieve a **consistent dataset**

Manipulating raw data – how to do it ethically?

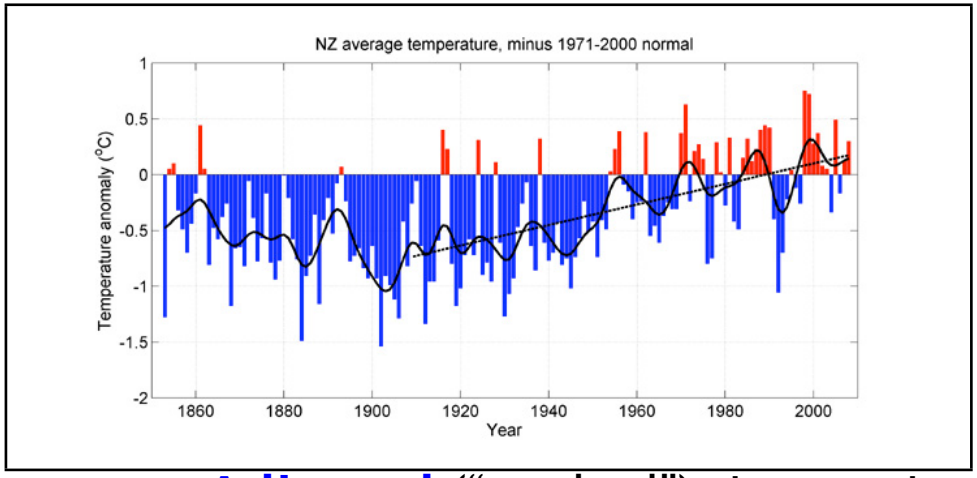
Manipulating data is **permitted** as long as:

- A **clear and consistent** methodology is applied **objectively**
  - can't "pick and choose"
  - **all** data items with a defined characteristic must be adjusted in the **same way**
- The adjustment methodology is **documented** as part of the research
- The **original data** are available for inspection.

Example: assessing climate change



Raw time series



Adjusted ("cooked") time series

Graph source: New Zealand Climate Science Coalition[1], climate change skeptics

Data source: National Institute of Water & Atmospheric Research (NIWA)[2]

[1] http://nzclimatescience.net/
[2] http://www.niwa.co.nz/

Plagiarism

Plagiarism: Knowingly representing the work of others as one's own

Several forms:

1. **Copying** someone else's work;
2. **Paraphrasing** someone else's work, i.e. saying the same thing with slightly different words and phrasing;
3. **Reporting** someone else's work (e.g. fieldwork) as if it were your own;
4. Getting someone else to do your work for you (**'ghostwriting'**);
5. (Using a particularly apt **term or phrase** which you didn't invent, without credit)

The wages of sin are …

BBC Mobile | News | Sport | Weather | Travel | TV

NEWS EUROPE

Home | UK | Africa | Asia-Pac | Europe | Latin America | Mid-East | South Asia | US & Canada | Business | He

1 March 2011 Last updated at 13:28 GMT

German Defence Minister Guttenberg resigns over thesis

German Defence Minister Karl-Theodor zu Guttenberg has stepped down after he was found to have copied large parts of his 2006 university doctorate thesis.

Mr Guttenberg, considered until recently a possible candidate for chancellor, has already been stripped of his PhD.

He told a news conference that it was "the most painful step of my life".

Mr Guttenberg said he was not able to defend himself any further

Tens of thousands of German academics have written to Chancellor Angela Merkel complaining about his conduct.

The plagiarism scandal led to him being nicknamed Baron Cut-and-Paste, Zu Copyberg and Zu Googleberg by the German media.

Related Stories

German minister loses doctorate

How to avoid plagiarizing

Three golden rules:

1. Everything you **write outside of quotation marks** must be the result of your **own creative effort**.

   Otherwise, you are taking credit for something you did not write.

2. Every **idea** that is not your own must be **credited** to the person(s) who conceived it.

   Otherwise you are taking credit for the other person's idea.

3. Every **fact** that you did not yourself establish must be **credited**. Otherwise you are claiming direct knowledge that you do not have. This includes field or lab. work actually done by others which you are reporting.

Example of plagiarism by copying

From a published book (1996) by Bergsma:

> Soil conservation is defined as the use of land, within the limits of economic practicability, according to its capabilities and the need to keep it permanently productive.

From an ITC MSc thesis script (mid 2000's):

> Soil conservation is defined as the use of land, within the limits of economic practicability, according to its capabilities and the need to keep it permanently productive.

Adding the citation is not enough

> Soil conservation is defined as the use of land, within the limits of economic practicability, according to its capabilities and the need to keep it permanently productive (Bergsma 1996).

This is not so bad, but it is **still plagiarism**. The author has credited Bergsma with the **idea** of this definition of soil conservation, but still implies that the actual **words** used are the author's interpretation, which they are not.

Correct but not elegant: verbatim quote with source:

> "Soil conservation is defined as the use of land, within the limits of economic practicability, according to its capabilities and the need to keep it permanently productive" (Bergsma 1996).

(continued . . .)

Putting the quote in context

It is more elegant, and more creative, to put the **relevant** part of the quote in quotation marks, and place that in your own context:

> Bergsma (1996) defines soil conservation as “the use of land, within the limits of economic practicability, according to its capabilities and the need to keep it permanently productive”.

or

> Soil conservation is defined by Bergsma (1996) as “the use of land, within the limits of economic practicability, according to its capabilities and the need to keep it permanently productive”.

▷ Bergsma, E. 1996. *Terminology for soil erosion and conservation*. Wageningen; Enschede: International Society of Soil Science; International Soil Reference and Information Centre (ISRIC); ITC

But did you need to quote anyway?

In our own words, we make the relevant points for our argument:

> Bergsma (1996) emphasizes three aspects of soil conservation: (1) using land according to its capability, (2) sustained production, and (3) economic feasibility. The present work is mainly concerned with the third aspect . . .

or synthesizing with other work, e.g.

> The concept of soil conservation was originally aimed at the physical protection of the soil from erosion at any cost and for indefinite time (Hudson 1981), but the emphasis is now on measures that are economically practicable and in line with the land's capabilities to provide productive and ecological services (Bergsma 1996).

Plagiarizing from digital source

(PDF, web pages, other people's documents . . .)

- Easy to **do**
- Easy to **catch**
- Easy to **prove and convict** (web search, plagiarism detection software)

Don't insult our intelligence by trying this. Save us both trouble and embarassment (more for you than for us).

Thank you.

Instead of plagiarizing ...

- **Think** and **write** for yourself!
  - It's **your** project, and you should want to express **your** ideas.
- **Summarize** one or several works in your **own words**
- **Quote** when you really need to use the text (e.g. to discuss another author's statement)
- Make a reference to the source and **do not include** in the text at all

Topic 3: Authorship

One of the two main principles of scientific ethics is “credit for work performed”.

A main source of credit is (co-)**authorship** of original scientific work reported in scientific journals, book chapters, or conference proceedings

Who can and should be an author?

Consult "Guide for Authors" for the target journal. For example, Elsevier[3]:

> "Authorship should be limited to those who have made a **significant contribution** to the **conception, design, execution, or interpretation** of the reported study. All those who have made significant contributions should be listed as co-authors. Where there are others who have participated in **certain substantive aspects** of the research project, they should be **acknowledged** or listed as contributors.
>
> The **corresponding author** should ensure that all appropriate co-authors and no inappropriate co-authors are included on the paper, and that **all co-authors have seen and approved** the final version of the paper and have agreed to its submission for publication."

[3] http://www.elsevier.com/wps/find/intro.cws_home/publishing

UT/ITC guidelines

- Published by the ITC library[4]
- Apply to publications originating from an MSc project connected to PhD research, where the PhD student acts as “MSc advisor”
- Also apply to MSc research that is incorporated into publications written by ITC staff, normally by the MSc supervisor

[4] http://www.itc.nl/library/copyrightguide.aspx

Five situations

1. The advisor or supervisor uses **minor material** from an MSc thesis, e.g. a graph, table, or quote. A literature reference to the MSc thesis is required at the point where the material is used (i.e. normal citation practice)

2. The article written by the advisor or supervisor has one or more **sections** that can be **directly traced** to material from an MSc thesis. The MSc student is co-author.

3. **Several related** MSc studies are included in a paper. The supervisor or advisor assembles and is the first author, the MSc students are all co-authors.

4. The material directly traceable to the MSc study makes up **more than half of the paper**. The MSc student takes the lead in authorship and is the first author, the advisor or supervisor is a co-author.

5. The MSc study is **substantially reworked** by the advisor or supervisor. These may be the first authors, and the MSc student a co-author.

Example: Minor material from an MSc thesis is used

- Bergsma, E., & Farshad, A. (2007). Monitoring erosion using microtopographic features. In J. de Graaff, J. Cameron, Sambran Sombatpanit, C. Pieri & J. Woodhill (Eds.), Monitoring and evaluation of soil conservation and watershed development projects (pp. 249-266). Enfield, UK: Science Publishers.

This article includes the statement:

"The sites were comparable in rainfall erosivity, general topography and soil (Table 4, basic data from Woldu, 1998)."

The cited thesis is:

- Hagos Dory Woldu. (1998). Assessment of the effect of present land use on soil degradation : a case study in Lom Kao area, central Thailand. Unpublished MSc, ITC, Enschede.

Example: A substantial section, but less than half, from the work of an MSc student; the supervisor or advisor is the first author

- Carranza, E. J. M., Hendro Wibowo, Barritt, S. D., & Prihadi Sumintadireja. (2008). Spatial data analysis and integration for regional - scale geothermal potential mapping, West Java, Indonesia. Geothermics, 37(3), 267-299.

This uses material from:

- Wibowo, H. (2006). Spatial data analysis and integration for regional scale geothermal prospectivity mapping, West Java, Indonesia. Unpublished MSc, ITC, Enschede.

Note the involvement in the article of additional authors from ITC (Barritt) and a research collaborator (Prihadi).

Example: The work of several MSc students is synthesized by the supervisor or advisor

- van Gils, H. A. M. J., Batsukh, O., Rossiter, D. G., Munthali, W., & Liberatoscioli, E. (2008). Forecasting the pattern and pace of Fagus forest expansion in Majella national park, Italy. Applied vegetation science, 11(4), 539-546.

This has material from two MSc theses:

- Batsukh, O. (2007). Beech forest expansion : spatial environmental modelling for prediction, Majella national park, 1975-2003 Italy. Unpublished MSc, ITC, Enschede.
- Munthali, W. (2006). Beech expansion : patterns, process and prediction. Unpublished MSc, ITC, Enschede.

These were both supervised by the same ITC staff; there is also a contributions from a collaborator in the fieldwork area.

Example: Mainly from one MSc thesis, written with the supervisor and/or advisor as co-author(s)

- Hengl, T., and Rossiter, D. G. (2003). Supervised landform classification to enhance and replace photo-interpretation in semi-detailed soil survey. Soil Science Society of America Journal, 67(6), 1810-1822.

This is based on:

- Hengl, T. (2000). Improving soil survey methodology using advanced mapping techniques and grid based modelling : case study, Baranja, Croatia. Unpublished MSc, ITC, Enschede.

Example: MSc study is substantially reworked in concept by the supervisor or advisor

- van Gils, H. A. M. J., & Loza Armand Ugon, A. V. (2006). What Drives Conversion of Tropical Forest in Carrasco Province, Bolivia? Ambio, 35(2), 81-85.

This is based on:

- Loza Armand Ugon, A. V. (2004). Spatial logistic model for tropical forest conversion : a case study of Carrasco province, 1986 - 2002, Bolivia. Unpublished MSc, ITC, Enschede.

Topic 4: Intellectual property and fair use

The intellectual, **intangible product** of a **creative effort**, such as writing, music, or a computer program, is as much the property of the creator as is a tangible object such as a work of art or a machine.

In some cases intellectual property is put into the public domain for free use, in other cases its **use is restricted**.

Copyright

Copyright (indicated by the © symbol) is the means by which an author asserts **ownership** of a work.

Laws vary between countries, and there are international treaties.

Basic idea: **the work belongs to the author**, who grants you certain **use rights**.

If you obtain the work legally, you can use it for your own purposes (e.g. read it for pleasure or instruction). Other uses are made explicit, for example:

> "**All rights reserved**. No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording or by any information storage or retrieval system, without the written permission of the publisher, except for brief passages quoted by a reviewer."
>
> – from Strang, G. 1986. *Introduction to applied mathematics*. Wellesley, MA: Wellesley-Cambridge Press

Fair use

To **compare** our work with that of others, may need to **quote** from the other work.

This sort of use is recognised by copyright law as **fair use**: use of the work for professional purposes; the quoting is necessary for the purpose.

Example: a book reviewer quoting passages from the reviewed book

Photocopies of printed matter, **printouts** of digital documents: 'reasonable' excerpts are considered fair use, otherwise buy the book or article

License agreements

Some materials are made available only under the terms of a **license agreement**.

The person who obtains it in a legal manner must first agree how it may be used.

This is very common with **computer programs** and **digital data**.

ITC: license from publishers for PDF copies of journal articles:

- to allow on-line access to the full text of journal articles
- license allows full use **within ITC**
- it is **forbidden to supply a third party** with the file; they would need to obtain it under their own license.

You, your employer, or your educational institution (ITC) is **liable for your actions**.

Restrictions on Datasets

Some digital data is supplied completely without restriction on what you can do with it, in particular data produced by the United States government.

Most digitial data is supplied along with an **end-user license agreement** (EULA), to which you must agree.

You can not use data in your thesis which is not legally yours to use:

- via your own license
- via ITC
- via some organisation of which you are considered part for licensing purposes.

Topic 5: Professionalism

Professionalism: scientists' role in the **society** in which they work, as representatives of a profession:

Carrying out professional activities correctly, according to **standards**, and within **societal norms**.

"Professional" vs. "amateur"
"official" vs. "individual"

Professional societies

- Many professional groups have “professional codes of ethics” or “standards of professional conduct”
- Include ethical standards **within the profession**
  - e.g. fair dealing with other professionals
- Also deal with how the professional should act within the **society at large**.
  - e.g. relations with government or business
- May have **legal standing**

Netherlands code of conduct

"Code of Conduct for Scientific Practice" from the association of universities (including ITC)

http://www.vsnu.nl/web/show/id=88938/langid=42

All researchers in the Netherlands (including ITC students) must follow this.

It is vague and open to interpretation, but the principles are clear.

5 points of the NCC

1. **Scrupulousness**

   Scientific activities are performed diligently, with care, resisting pressure to cut corners in order to achieve

2. **Reliability**

   The scientist makes every effort for their work to be accurate and thorough, thus reliable.

3. **Verifiability**:

   Any publication based on research must clearly state the basis for the data and conclusions, including the data source and analysis methods; all of this so that the reader can in principle independently verify the work.

(continued . . .)

NCC (continued)

4. **Impartiality**

   In scientific activities, the scientist must have no other interest than science, and be prepared to prove this. This is most relevant when the scientist works for industry or has commercial interests.

5. **Independence**

   Scientists operate in a context of academic freedom and independence from interference. If this is not possible for commercial, political or institutional reasons, this must be clearly stated and justified.

Topic 6: The social responsibility of the scientist

Science is now big business and an integral part of society.

It has a large effect on society (“age of science”).

The scientist can not pretend to be “value neutral”, choices must be made.

Selection of a research topic

Important ethical decisions are made at the beginning of a research project, with the selection of a research topic.

- Would the results of the research be **useful** to society?
- Is the topic related to a **social problem** of importance?
- Would the results of the research be **socially valuable**, or at least not damaging?
- Are various sectors of society **marginalized** or even directly **harmed** by the research?

Examples of ethical dilemmas – selecting a topic

Many research topics pose ethical problems, for example:

- Any **remote-sensing** project by its nature (view from above) invades the privacy of individual land owners; it also violates the sovereignity of the country imaged.
- Any **natural resources survey** or land suitability evaluation project implies that knowledge of these will be given to people outside the affected area, who may make planning, investment or migration decisions that may not benefit the local population.
- A design thesis that builds on a specific **computer program** is implicitly endorsing that program and, if it is a **commercial program**, promoting the financial interests of the company that produced it (ESRI, Microsoft . . .).

  Conversely, use of an **open-source** program may reduce commercial opportunities but increase the overall productivity of the research community.

Trendiness and 'political correctness'

- The modern research establishment runs on public (government) or charitable (foundations) **funding**.
- These have explicit **agendas for research**.
- Agendas are based on explicit **social goals**, which are translated into research priorities.
  - Example: European Commission's (EC) research frameworks, currently in the seventh round ("FP7")
- Researchers often have to **tailor their projects** to the demands (and jargon) of the funding agencies

"He who pays the piper calls the tune"

Example: EC FP7 "Food, Agriculture and Fisheries" theme

"The advancement of knowledge in the **sustainable management**, production and use of biological resources (microbial, plant and animal) will provide the basis for **safer**, **eco-efficient** and **competitive** products and services for agriculture, fisheries, feed, food, health, forest-based and related industries.

Important contributions to the implementation of existing and prospective **policies and regulations** in the area of public, animal and plant health and consumer protection are anticipated.

New **renewable energy sources** will be supported under the concept of a European **knowledge-based bio-economy**."

. . . whatever a "knowledge-based bio-economy" may be !

Political correctness

- Avoiding language or implications that might be considered **offensive** by some group with a self-identity
  - Examples: national, ethnic, gender, age, social status.
- **"Hot button" topics**, not to be mentioned (or even thought about!)
- Implies that some ideas are not acceptable for research, or even for discussion.
- But ignoring reality won't make it go away, there may be real problems which research can investigate
- Solution: avoid offense, support statements with **evidence**, be aware of own biases (see below)

The scientist as a social animal

Scientists are humans, so have:

- **values**
- **biases**
- **subjectivity**

The scientist should **recognize** these and **account for them** in research activities.

Interactions with colleagues

- Governed by **narrowly-defined scientific ethics**, particularly the rules for assigning credit for work performed.

- However, there are often **cultural** differences (both general and scientific) in:

  * working methods
  * expectations of roles and responsibilities
  * priorities
  * attitudes towards authorities
  * communication style

  which can hinder scientific progress.

- **Economic** and **status** differences can exacerbate these cultural differences

- **Solutions**: awareness, sensitivity, communication, flexibility and common sense

Local populations

- Humans (researchers) studying other humans (local population)
- Inherent **differences** in status, economic power, priorities
- What is the proper **relation** between researcher and subject?
- Is there always a **win-win** solution (both parties benefit) or is there always an element of **exploitation**?

Classic example: native knowledge of plants → commercial drugs

Examples of ethical dilemmas

- How should local people be approached? What information about the research purpose should be given?
- Will the results of the research be 'returned', and if so, in what form?
- What to do if the research is not in the benefit, or even to the detriment, of local populations? Example: studying soil erosion vs. farming practices, this may lead to a ban on certain crops or management on certain lands (e.g. steep slopes), which is a short-term economic loss to the farmers?
- If surveys are to be performed, what information about them is given to the participants? Should they be paid or otherwise rewarded? (continued . . .)

Dilemmas (continued)

- What are ethical methods of asking questions or making observations? Can subjects be “tricked” with false promises or pretexts?
- How intimate should the researcher be with the population? Does the researcher sacrifice neutrality or objectivity by identifying too closely with the subjects or target group?
- How should researchers balance their own cultural values with those of their subjects?
- How to extract reliable information within cultural limitations? Example: It is considered improper in the local context for a male researcher to talk directly with a female subject; should the researcher trust a male relative’s interepretation of what the female says?

Who benefits from your research project?

Ethics in the widest sense: Who benefits from research?

1. **You**: advances your career; able to do interesting work; allows you to satisfy curiosity, feed your ego . . .

2. Your **family**: income, status

3. **ITC**: receive credit, show ability to train students, attract students and support;

4. Your **sponsor** (home organization, funding agency): they get what they paid for;

5. The **scientific enterprise** in general: more is known;

6. Future **employers**;

(continued . . .)

Who benefits? (2)

Do these benefit? How? Can you be sure?

1. **Society** as a whole?

2. The **individuals or communities** who helped you or made your research possible?

Can your work be used for **harm**?

- "Knowledge is power"; will the 'wrong' people have knowledge that helps them?

- Repressive governments, exploitation . . .

"Life is complicated . . ."