 # Link of Lecture 6 #

! see also http://privatewww.essex.ac.uk/~scholp/litrevsarc.htm

^^ وضع هذا الرابط للاستزاده .. في النقاط التي يجب مراعاتها خلال كتابة الـ LR

The tongue-in-cheek guide to writing a literature review as part of an empirical research project

Include everything faintly connected with your topic. Don't bother to sift out what is central and omit material that is distant: you don't want to be short of material!

A 'random walk' through a topic is so much more interesting than a structured, logical progression with lots of headings and subheadings for different aspects of the topic. The reader likes a mystery tour in a piece of academic writing. If it is all vaguely to do with your research topic, what more can the reader expect?

Especially, don't tell the reader beforehand what areas you are going to review, and why.

A nice idea is to use the title of a chapter or section in a review just as a starting point. Then take the reader off into a mystery tour of all sorts of areas that don't belong under that heading.

Make sure the review is so broad and long that there is no room for anything much original of your own. Quoting other people is so much more impressive than your own comment or analysis, or links shown with your own experience, country, project etc.

Give all your sources equal weight. If it is published somewhere it must be true and all truth is equal, yes? That means there is

No need to check where your source got their information from: whether they are just quoting someone else or actually did original research themselves or indeed if it is just a personal opinion.

No need to criticise the reasoning used by any source to arrive at a statement. No need to be bothered about whether your source is consistent with current relevant theories in the field.

No need to bother with trivia like whether their research method was sound or not, whether their questionnaire questions were ambiguous, what subjects they had etc.., or whether they are just retailing a personal anecdote. It's the ideas that count.

No need for you to compare what anyone says with what anyone else says and add any argument of your own as to which is more likely to be true.

If two sources are using the same terms for what they are talking about, then they must be talking about the same thing, right? After all, in applied linguistics and ELT people never vary in how they use key terminology. E.g. they all use 'communicative' for the same thing, they all mean the same thing by 'function' etc…. So you never need to question if they really are talking about the same thing as each other, or you…

If two people make the same point it must be right. Better if several say it, quoting each other.

There's joy in repetition. If you've made a point once, quoting someone's opinion on something or giving some fact, it must be worth doing again. In particular make sure you

Separate the repetition of the same point by a few pages so with luck the reader will think it is a new point

Even better, put it in a new section or chapter with a different title

Put it in different words, with a different source reference, and never mention that it is a point you have already made

Introduce it as a new point, even though it isn't.

When you are making a series of points from different sources, make sure you yourself never distinguish between where they are really saying the same thing and where they are saying the opposite. That is not your place. Just string it all together and leave the reader to figure it out.

If two sources clearly say different things on the same point, make sure you don't offend anyone by pointing this out. Above all don't add any reasoning of your own to choose between them.

It is much safer just to cite different opinions and never make it clear which you agree with and are going to adopt for your work and which not. After all, you might pick the wrong one.

The best way to be critical about someone's work is to cite what other people have said about it. No point in hearing your voice as well.

It's especially handy when sources use different terminology for much the same thing, as often happens in applied linguistics and ELT. Be sure not to point this out. E.g. an article about 'consolidation' or about 'mnemonics' must surely be about something different from 'retention'.

Also useful is to cite other people's research in as little detail as possible. Don't bother to mention what country it was in (the same one your project will be in or not?)**,** what languages involved, what level of learners or whatever. That way the account is so vague it looks as if it could apply to almost anything, including your research. After all, for example, what is said about teaching writing at one level in one particular teaching situation in one country must surely apply to any situation on any country, including the one your study is going to be on?

If you do do a longer review of a key article, be sure to follow the agenda of the article itself, even if it is different from yours. It would not do just to cherry-pick the points that are relevant to your own project and leave out the rest.

The main point to extract from a summary of an article – the 'importance' of the article - is what the author of it thought was important, not what is important about it for YOUR study.

Don't bother to summarise the overall picture that emerges from a group of sources you go over. After all, the reader should be made to do some of the dissertation work for you.

If you do provide a summary, make sure it is a summary of everything you reviewed, not just of the points derived from all that which are relevant to your own project.

Assuming you do go on to report some empirical work of your own after the review, make sure there is as little connection as possible with the review. After all, the two are quite different parts of the work. For example

In your review, never refer to the study you are going to do, or extract any predictions for what your study might find. Leave the reader to spot the connection later

Better, make your study deal with something different from what was covered in the literature review. You don't want the reader to get bored

If you do comment on your sources, be sure to point out the interest and importance of issues, variables etc. that in fact you are not going to include in your own study. The reader will enjoy the surprise of having been led to expect that you are going to gather data on one thing and find later that you have actually gathered data on something quite different

It would be bad form to revise your lit review after gathering your data to make sure it connects. Once you have written it, leave it

If your own project has a list of research questions or hypotheses, never point out what bits of the literature review (if any) prompted them. Just list them and leave the reader to figure out what there was in the previous 50 pages of review that had any connection with them

Don't relate your 'method' to that used by other studies. You don't want to look unoriginal or appear to have learnt anything from others' experience or mistakes

If you are evaluating course materials from your country, make sure the criteria you use to evaluate them have nothing to do with the theories and research talked about in the literature review. They can't have any connection with your country, after all. Just dream up a miscellaneous set of your own

If you are administering a questionnaire the questions should be made up out of your head. Again, why learn from others' experience?

When you get the results, just summarise them. It would be presumptuous to try to relate them to any other research reported earlier in your review.

PJS Written in MA dissertation and PhD thesis shock, Oct. 96 with slight additions 04

PS Just in case you have not spotted it, the above is SARCASTIC.

A good review does the opposite of all those things.

 # Link of Lecture 9 #

 # RH & RQ محتوى الرابط الخاص بـ

http://privatewww.essex.ac.uk/~scholp/Hypotheses05.htm

Formulating questions and hypotheses about differences and relationships

The key to the success of an entire classical investigation is to state clearly the research questions and/or hypotheses (RQs and RHs) for the specific topic of interest. As we have seen, they must be 'sourced' or 'founded' in appropriate background. They should also be RQs and RHs you are interested in. And later, when you get Results you must not forget to refer back to them! Above all they must be empirically researchable - i.e. have the capability for a Method to be found to follow them up. That means not too many, and not requiring subjects that in practice you could never get hold of, and not worded in vague terms so one cannot see pretty directly what subjects are to be involved and what variables measured etc… E.g. *Does mother tongue affect second language acquisition?* is far too vague. You need to word it in terms of at least what languages you mean, what specific learners, what aspect of language will have its acquisition looked at, etc.

Now, though it may appear somewhat unnecessary and pedantic at this stage, it will prove invaluable for later analysis of results and the use of statistics to think carefully from now exactly how these RQs or RHs are stated, and if necessary reword them to suit what you really do want to investigate, in one of a limited number of general wordings. A number of things need attention:

(a) Research question or research hypothesis?

Are you really asking a question to which you have no idea of the answer, or have you really got an expectation of the answer (i.e. a hypothesis)? There is a difference between specifying your aim as

1. to answer the specific question:

   Is there a difference between Taiwanese High School girls and boys in attitudes to learning English?

   (or some equivalent wording such as: I want to see if there is a difference between Taiwanese High School girls and boys in attitudes to learning English; I am interested in Taiwanese High School girls' and boys' attitudes to learning English)

2. to test/prove the hypothesis that:

   There is a difference between Taiwanese High School girls and boys in attitudes to learning English.

   (or some equivalent wording such as: Taiwanese High School girls and boys do not have the same attitudes to learning English)

If you really have no expectation, (1) is the appropriate formulation. Your research is 'exploratory'. But in fact nine times out of ten when you formulate your aim as a question, you actually have an expectation of the answer (based on previous research etc.), otherwise you would not have asked that specific question, so you are into 'hypothesis-testing' or 'confirmatory' research (2). It is then good practice to reformulate your question overtly as a hypothesis - i.e. a statement of what you expect. A piece of classical research usually sounds better if its essential plot is "I thought X might

be the case (for W reason), I used Y method to see if it was, and hey presto the results Z show it was (or wasn't)".

Reasons for stating a RQ rather than a RH include:

Past research gives conflicting results

Current theories suggest opposite conclusions

There genuinely is no past research or theory bearing on this, even remotely

You have simply omitted to find or read the relevant background literature!

An example of conflicting expectations arose recently with a PhD student researching the relationship between extraversion and use of vocabulary learning strategies (VLS). According to past research (in psychology) extraverts have better short term memory than introverts, but introverts have better long term memory. What would that imply for VLS? If introverts have better long term memory, would they therefore use strategies that aid LTM (e.g. various word association strategies) more than extraverts? Or, because they have good LTM in any case, would they use them less because they have less need of them? Here one can pose an interesting RQ, but not be sure which of two possible RHs to go for in advance of data gathering.

No-no! Hypotheses are supposed to be formulated before you gather data. It is not permitted in the Classical approach to gather data first, then write some hypotheses which predict that you will find just what you actually have found! The most that is permitted is that, after gathering the data, you may realise that there are some things you gathered data about that you should have had hypotheses for, but forgot to formulate them and write them down earlier. Then you can go back and put some hypotheses in place, but you should not of course word them to predict what you have since found in your data, only on what the literature etc. says that existed before you gathered your data.

Side note. As we have seen, in contrast with the Classical approach, the Ethnographic research method involves (rather general) RQs rather than RHs, and these may not even be decided in advance, as in the Classical model, but as the research goes along. In my view it is often a good idea to add an element of this in a classical project by including in the Method opportunities to gather data not just related to the prestated RQs and RHs, but with the potential to 'find' new RQs. That means having open as well as closed questions in questionnaires, having informal interviews with subjects where they can talk about anything that comes to mind, carefully observing what happens while subjects do your test/task or whatever with an eye to spotting something 'unexpected' happening, and so on.

(b) What are the 'null' and 'alternative' hypotheses?

Technically hypotheses come in two forms - one the negative of the other. If there is an expectation that something will prove to be the case, there is an expectation that its opposite will not prove to be the case. It may seem a bit pedantic at this point to talk in both these terms, because supporting one by definition means not supporting the other, but it will turn out later (when doing statistics) to be useful to formulate both expectations. E.g.

There is a difference between Taiwanese High School girls and boys in attitudes to learning English.

There is no difference between Taiwanese High School girls and boys in attitudes to learning English

As a matter of technical terminology, one of these hypotheses is usually referred to as the 'null hypothesis' or 'H_0' or, in some statistical discussions as 'E'. The other is the 'alternative hypothesis' or 'H_1'. The null hypothesis is the negative form that refers to "no difference" or "no relationship" or "non-existence" or the like; the alternative hypothesis is the one that claims a difference etc. Note that this distinction does not coincide exactly with the distinction between the hypothesis you expect will be confirmed and the one you think will not. Usually the alternative hypothesis is the one you believe in and the null hypothesis is the one that you think is not true, but that is not always so. (Note, when it is used, the label 'E' stands for 'expected'. But that means 'What you would expect under the null hypothesis', not 'The hypothesis that the researcher expects to be confirmed').

A research question where you have no expectation is tantamount to stating both null and alternative hypotheses, but not choosing to support either in advance of gathering any data.

With more complicated hypotheses there may be an H_2 etc. as well. And in research with many variables there may be many hypotheses, or a mixture of RHs and RQs.

People vary in how far they actually state all the null and alternative hypotheses in a piece of research. But they need to be there conceptually. Probably it is most useful to state the ones one actually is predicting (whether null or alternative) and leave the other partner unstated.

One last point. Though one's interest may be mainly in 'proving the alternative hypothesis', it turns out that by classical research one cannot so much prove that hypothesis as disprove the corresponding null hypothesis. Statistical tests (those things you may have heard of like chi squared, t tests, ANOVA,...) are mostly geared to helping you judge if null hypotheses can be rejected or not, on the evidence of the data you have gathered. Thus one gets into a fair amount of confusing thinking in double negatives!

(c) What are the key variables in the question/hypothesis, and is a comparison of some sort involved?

Classical research is basically about variables - i.e. aspects of people, situations, language etc. (i.e of cases) that can take on different values or levels. Social class or intelligence of people, people's language proficiency measured by IELTS, degree of formality of situations, part of speech of words, type of teaching method, are all variables. Classical RHs and RQs can then be seen as questions or statements about variables - very often the interrelationship between more than one variable. These variables have to be clearly identified in the wording, because when you get to the Method part of the research you then have the job of actually identifying/ measuring/creating these variables, so you need to be clear what exactly they are.

In the examples so far, the variables have been fairly clearly two - sex (male versus female) and attitude to learning English - and a comparison was involved. English is apparently intended to be a constant, though one could of course formulate hypotheses that also compared languages (English with Japanese for example). Taiwanese High School students are also a constant in the project as stated so far, though again one could introduce comparisons with the attitudes of learners at other levels or in other countries.

Two-variable questions/hypotheses with some comparison involved are very common. However, so you can see the difference, I give a few examples here where RQs/RHs are not straightforward

comparisons of some sort follow (though actually they probably could be reworded as some sort of comparison ultimately).

i) Rarely, RHs and RQs are simply about the 'existence' or occurrence of something, not an obvious comparison. I.e. they relate to a variable with two values - something happens or it doesn't. E.g.

H_0 No natural language has more than 65 consonant phonemes.

is such a hypothesis. It is disproved simply by finding a language with more than that number of consonants.

ii) Now, what about this RQ?

Do Taiwan primary school teachers agree that an excellent way to teach English is communicatively?

This could mean that you just intend to ask these teachers about the communicative approach and see if the majority say yes to it or give it a high rating for suitability. No obvious comparison: you are interested only in the results for this one group on one variable. Instead it requires reference to some absolute scale of amount of agreement about the excellence of this method and asks if these teachers will score near the top end or not. (Or… there could be implied comparisons with how they would rate other methods or what other teachers would say, in which instance the RQ should be reworded… see (e) below).

iii) Some RQs in purely 'exploratory' research differ from the common classical type in being more elementary - simply about what categories might be relevant to describe some phenomenon (a qualitative matter, which needs to be settled prior to making comparisons or indeed counting anything). They are of course more typical of the Ethnographic type of research. E.g.

What listening strategies do intermediate Hong Kong learners of English use?

In classical terms this could mean that the project is to establish a new variable called 'types of listening strategy' and work out from data what suitable categories it should contain. Or of course it could involve using a pre-existing classification of strategies, from someone else's study, or modifying it.... See also further (e) below for another way of interpreting this RQ, as involving a hidden comparison.

(d) Is the research on differences or relationships, or formulatable either way?

It will turn out to be useful to get used to thinking of what one might vaguely call "comparisons" in two ways. Are they stateable basically as a matter of 'difference' or 'relationship' or both?

Often (but not always) they can be thought of either way. Earlier we had this RH worded as a matter of difference:

H_1: There is a difference between Taiwanese High School girls and boys in attitudes to learning English.

Note how the difference is stated between two or more categories or values (e.g. ‘girls’ – ‘boys’ or 'male' - 'female') which together constitute a variable (here 'sex'). The way in which they supposedly differ (attitude) is a second variable.

Reworded as a matter of relationship:

H_1: There is a relationship between sex and attitude to learning English among Taiwanese High School students.

A relationship must be stated between variables (here two). It makes no sense to produce a mixed formulation of the type:

> H_1: There is a relationship between girls and boys in attitudes to learning English.

However, not all RHs and RQs can be easily restated in either form.

> There are more women than men teaching English in Taiwanese High Schools.

predicts a difference of a sort, but no relationship, because there is only one variable involved: sex of teacher.

Similarly:

> Is this class doing worse in reading than the average for learners in general of the same year?

Again this is a difference question that cannot readily be translated into a relationship one.

Conversely the following concerns a relationship that is not exactly expressible as a matter of difference:

> The speed with which people can access a word from their mental lexicon is related to the word's frequency in the language

It is, however, usually possible to turn an example like this into a difference hypothesis, at the cost of changing it slightly, and losing some precision. Instead of talking about a continuous variable 'frequency of words', as implied in the wording above, you could change it to 'frequency bands of words' (e.g. high frequency words - mid frequency ones - low frequency ones). Then you could have a difference formulation of the type

> There is a difference between high and low frequency words in the speed with which they are accessed from ML.

Note on the word 'relationship'. This is the most useful general word for what we are talking about here. Do not use 'association' or 'correlation' as synonyms of 'relationship'. Those two label specific types of relationship defined in statistics, so should only be used where appropriate (In summary, the relationship measured between sets of scores by the Pearson r or Spearman rho coefficient is a 'correlation'; the relationship measured between categories with chi squared or the like is 'association'). Another pitfall is to interpret all relationships as causal. There is a difference between talking about the relationship between x and y and referring to the effect of x on y or saying that x causes/leads to/affects/influences y. We will find that statistics is good at telling you where there are differences or relationships, but not so good at saying where they are causal or not. Often one may think that a relationship is causal, but it can be hard to prove more than a relationship, because proof of cause typically needs an experimental research design, and in many areas of language study one cannot readily do experiments. For instance, we may show that sex is related to attitude to RP, but does that prove that sex itself is the cause? It wouldn't seem likely that the difference would be part of biological sex, anyway. It is quite possible that the real cause is something else that happens to be related to sex. For instance, it is known that teachers in school, and parents at home (dependent to a degree on social class) tend to talk to girls differently from boys and project different attitudes. Maybe then it is these people who really 'cause' the different attitudes, and sex is a sort of intermediary only.

(e) Is there an incomplete or hidden or forgotten comparison?

A great source of lack of clarity about research is the question or hypothesis formulated as an incomplete comparison. E.g.

Do Taiwanese High School boys have a less favourable attitude to learning English?

sounds OK until one thinks about it in detail. Is the question about whether Taiwanese High School boys have a less favourable attitude than Taiwanese High School girls or than Junior School boys? Or is the claim that Taiwanese High School boys have a less favourable attitude to learning English than to learning other languages? It is always advisable to complete the than part of any hypothesis or question containing a comparison.

Similar are questions or hypotheses with a 'hidden comparison', implied but not stated clearly.

Women use a lot of question tags

appears to be just a claim about women. But on closer inspection there is surely an implied comparison. How can one assess what "a lot" really is without comparing with some other sort of person (e.g. men), or the frequency of some other linguistic unit (e.g. questions without tags)? One would not prove much just by examining women's speech alone and counting question tags. An example of a better formulated hypothesis would be:

H_1: Women use more question tags than men do.

Yet another way in which comparisons can get disguised or forgotten is where part of the comparison actually would be 'filled in' from previous findings by other researchers, not something be covered directly in the present study.

What listening strategies do learners of English in Hong Kong use?

could mean several things. It appears to announce a research project just in Hong Kong - i.e. in the jargon place is not a 'variable' but a 'constant'. It could indeed mean that (see end of (c) above). However, you might feel it really implies a comparison with other places:

Do learners of English in Hong Kong use different listening strategies from learners elsewhere?

This could be studied by using information about learners in other places, as reported in studies by other people. Certainly, if other such research exists (which it does in this instance), it would be inadvisable to disregard it and not make some such comparison at some point.

In short, it is well worth considering if there is not a comparison that could be usefully made, or indeed is implied but not stated, in your hypothesis, and build it in overtly.

(f) What is the exact focus of the research and what is being assumed?

Your research question or hypothesis must make absolutely clear what the focus of the research is - the claim to be tested or the matter in question. Distinguish that from what is being assumed to be true, since any question/hypothesis normally assumes that some other questions have already been answered (or that the answers are self-evident), or assumes other prior hypotheses to be already proved (or so obvious as to be able to be assumed proved). The trick is to make sure that what you are focussing on is clear and unambiguous, and what you are assuming can reasonably be assumed and is not something that should really be established itself first.

For instance the hypothesis (H_1):

> There is a difference between Taiwanese High School girls and boys in attitudes to learning English

is reasonably clear in focus and assumes fairly little. It focusses on the idea of there being some difference between males and females in the attitudes (without predicting which sort of person has the more favourable attitude or how much exactly - hence the label 'nondirectional' for this sort of hypothesis). It assumes that girls and boys do have discoverable attitudes to learning English - a reasonable presupposition at High School age - and just needs a more precise characterisation of what attitudes are involved, what grade in High School, and the like.

However, if past research had already established the above hypothesis to be correct, it would be assuming too little, unless the current research was intended merely as a 'replication', confirming what someone else found. A stronger claim would be needed for this, making it a followup piece of research. One wants to avoid setting up 'straw man' hypotheses or the response may be "So what? We knew that already". There are two common ways to arrive at stronger/more elaborate hypotheses.

i) As a followup here maybe (depending on the literature) one would introduce another variable. That can often produce an interesting new project out of one that would otherwise have gone over well-trodden territory:

> H_1: There is a difference between Taiwanese High School girls and boys in attitudes to learning English, but it differs depending on whether the school is in a city or in the country.

In fact as soon as you have three variables in the frame, there are potentially three hypotheses in all that might be worth formulating (some perhaps already supported by earlier work): the above, which predicts an 'interaction' between sex and location of school in the effect on students' attitude, and the following two 'simple effect' or 'main effect' hypotheses:

> H_2: There is a difference between Taiwanese High School girls' and boys' attitudes to learning English (regardless of location of school).
>
> H_3: There is a difference between students in Taiwanese city High Schools and those in country High Schools in attitudes to learning English (regardless of sex).

ii) Another common way of elaborating a simple hypothesis is to make it 'directional':

> Taiwanese High School girls have a more favourable attitude to learning English than do Taiwanese High School boys.

Or, in strict difference wording:

> There is a difference between Taiwanese High School girls and boys in attitudes to learning English, specifically girls have a more favourable attitude than boys.

Or, in strict relationship wording:

> There is a relationship between sex and attitude to learning English among Taiwanese High School students: specifically girls have a more favourable attitude than boys.

This clearly is a stronger hypothesis than the original one, as it does not just hypothesise a difference, but also which group will score more highly.

In the above, one is in effect considering two possible results: Alternative H: girls score higher for favourable attitude than boys do; Null H: any other result… i.e. boys score higher than girls, or both the same. An even stronger form of this directional hypothesis would be where the possibility of boys having the more favourable attitude is assumed to have been ruled out as altogether impossible. The two outcomes one is considering then are: Alternative H: girls score higher for favourable attitude than boys do; Null H: both girls and boys score the same. This is called the 'one-tailed' hypothesis, but researchers in social sciences, language study etc. rarely feel confident enough to be able to make the necessary assumptions, so routinely stick with statistical tests designed for the 'two-tail' hypothesis. (The reason for talking about 'tails' will appear when we get to the statistics - they are the tails of a curve shape on a graph of a certain sort).

Often in the relationship wording the words 'positive' and 'negative' carry the information on direction:

There is a relationship between learner level and use of third person -*s*

Is non-directional. Compare:

There is a positive relationship between learner level and use of third person -*s*

Note: Students often seem to misuse the words assume, assumption as if they referred to hypotheses or just argumentation. For example: 'The assumption that strategies developed in the native language transfer to the foreign/second language is supported partially by data in this investigation'. Here what is meant is 'The hypothesis that…..'. An assumption is more than a hypothesis: it is what you are prepared to take as given and does not need proving, rather than an expectation about what you want to prove. Sometimes I even read students talking about what is shown by the results of a study as being 'assumptions'.

(g) Other kinds of question and hypothesis are needed?

Most common RH forms one sees in classical research are either non-directional or directional, formulated in either difference or relationship language (or in wording that easily converts to that), and involve just two variables. These are the ones that the most familiar basic statistical tests are designed to deal with. However, more precise and complicated hypotheses are possible, usually with the assumption that simpler ones have been proved already.

Hypotheses or questions concerning more than three variables at once can be formulated, but may get complicated and confusing. E.g. a common kind of exploratory question in classical research is of the type:

> Out of instrumental motivation, integrative motivation, intelligence, field dependence, parental interest, and extroversion, which combination is the best predictor of language learning success?

This mentions seven variables, assumes all of the first six are related in some way to the last, but asks which of the first six collectively relate most closely to the last, in such a way that the remaining ones then have little influence.

A comparison may go beyond the types we have seen to involve questions or hypotheses about the actual size of a difference or the precise nature of a relationship (there are several types best understood from graphic representation). These will not be detailed here. E.g.

By how much do women exceed men in frequency of question tag use?

Note, that assumes not only that a difference exists between sexes, but also in a particular direction. It goes a further step to ask about the extent of the difference.

What is the precise relationship between social distance and politeness?

This assumes there is a relationship, but asks what kind: e.g. Are people successively more polite when talking to people who are more socially distant from them? Or are they less polite with very close and very distant interlocutors, more polite with intermediate cases?

(h) How many of these questions/hypotheses should there be?

Often people overlook just how many RHs and RQs are actually involved in their research. They just state one or two, but then in Method use a questionnaire with 50 questions, many of which imply possible further RHs and RQs that should have been formulated (with their attendant background review). With questionnaires you can rapidly get into a situation where you are unable either to deal properly with all the background and possible RHs and RQs presupposed by the questions or with the results that follow, because they cover so much in a short space. (Note a questionnaire question does not always correspond one to one with a research question, but often does). Generally it is a bad idea to go back after data gathering and fill in the RHs one should have thought of in advance, but it has been known to happen...

Thus some techniques used as part of one's research Method may involve a lot of RHs/RQs to a small amount of Method activity. Conversely one hypothesis, with only two variables involved, may entail a large amount of Method. E.g. the one RH

> Use of authentic materials in EFL increases learners' intrinsic motivation (more than the use of non-authentic materials does)

involved a student in several months in Korea teaching different classes with different materials and observing the effects.

Task: sharpen up the RQ or RH in clear difference or relationship wording

The core of a researchable classical research project is usually stated in the form of one or more hypotheses (and/or maybe research questions). Usually there is an expectation (from theory or past research) of what the result will be, or should be if a certain theory is correct (unless the research is purely exploratory). Try reformulating the projects below in terms of carefully stated hypotheses, making clear what the focus is and what the relevant variables are. Wherever possible, identify what is to be investigated in terms of differences and/or relationships, showing direction where relevant. I.e. attempt to use one of these wordings, or similar:

> There is a difference between value A of Variable 1 and value B of Variable 1 in Variable 2
> OR There is a difference in Variable 2 between value A of Variable 1 and value B of Variable 1
> OR A will score higher/lower than B on Variable 2
>
> There is a (positive/negative) relationship between Variable 1 and Variable 2
> OR 1 and 2 are related (positively/negatively).

E.g. *Compare a child's mean length of utterance (MLU) from speech samples obtained in different circumstances* might become *There is a difference between the MLU of a 5 year old child's speech when addressing its mother and that when addressing a stranger*.

(Here V1 = addressee, with value A = mother and value B = stranger; V2 = MLU)

But beware not to force things into this mould if it seems to entirely change the spirit of the project. Along the way you may need to sharpen up the statement of what variables, constants and people are involved, and where there are incomplete comparisons make educated guesses about how they might sensibly be completed. Give a few words of explanation about how/why you are making such refinements.

Student research/hypothesis proposals:

There is a relationship between students' reading comprehension in L2 and the level of linguistic difficulty of the two passages being read.

Integrated teaching with listening and speaking is more effective to promote students' communicative ability

Increasing level of proficiency in L2 does not avoid high interference of L1 into learning process of L2.

The students use dictionary more for translation into L1

There is a relationship between using reading strategies while reading in the foreign language (English) and using reading strategies while reading in the native language (Spanish).

The traditional way of designing reading comprehension exercises/questions as well as the teacher's way of using them cannot provide efficient help for learners to build up different aspects of reading skills and to recognise discourse features

The use of bilingual dictionaries involved an excessive amount of time and thus led to many errors caused by language interference

Teachers might not use CLT because of the students' passive attitude

Is there any difference between Pakistani Intermediate students' attitudes towards English language on the basis of their socio-economic status?

Is Instrumental motivation more applicable to the learning of English than Integrative motivation in the Pakistani setting?

Do students <learners of English in Saudi schools> forget words over vacation time?

What kinds of dictionaries are most popular among Saudi learners?

What are their <Saudi learners'> habits in using dictionaries?

 # Link of Lecture 13 #

One variable: graphs and descriptive statistics

WHEN DO YOU NEED THEM?

When does one ever want to look at results for just one variable? True, most classical research involves questions/hypotheses that entail looking at relationships between at least two variables. But here are some common situations where you want to look at graphs and descriptive statistics for cases on one variable:

● Some studies look directly at more or less all of an entire population of interest, and have questions about single variables, so the results just need to be given as graphs and descriptive statistics. E.g.
-- A census survey to find out how many people in Wales speak Welsh shows that 20.8% claim to.
-- A teacher has the feeling (hypothesis) that her vocab teaching (or the students' learning of vocab) is not very successful. She tests her class to see how much of the vocabulary of the last ten lessons they have learnt, as part of an action research project to improve her vocabulary teaching.

● In any study, however many variables are involved, it is often valuable to describe our cases in terms of each of a set of variables that are not central to the investigation, as part of your control of unwanted factors (cf CVs and SAMPLING). E.g.
-- You are interested in the opinions of Greek learners in private schools in Greece about their English course materials, so you send 50 questionnaires to your friend who works in one to distribute for you. You get thirty-one back. On the questionnaire, apart from questions about the central variables of your study, you might ask questions to elicit their first language, age, gender, experience of English outside school etc. You then check each of these separately to see if it suggests your sample is unusual in any way (?unexpectedly many girls), has odd cases in it (?two who say their first language is Bulgarian) etc. You might well display proportions and graphs for genders, age groups, respondents vs non-respondents etc. as part of your report on your subjects.

● In any study, however many variables are involved, you may be in the position of deciding groups of cases on the basis of information gathered about them, rather than in advance. E.g.
-- In the above example you might actually want to make up groups out of your subjects, using their questionnaire responses, to use as EVs. Examining the 'experience of English outside school' variable you find there are ten who have been abroad to English speaking countries, so you might set up a two category EV on this basis to see if opinions about course materials relate to having/not having this experience at all. If so, would you have any hypothesis about the answer? Anyway, you might display a graph or table showing the proportions.

● In any study, however many variables are involved, it is often valuable to look at results for all cases on each dependent variable or condition separately and/or in each group separately as well as doing what is necessary to establish relationships between variables. E.g.
-- In our example of gender and attitude to RP, with the hypothesis that there is at attitude difference between genders. Apart from doing the relevant two-variable graphs and statistics one would do well to explore the data with histograms for each group separately and the whole set of

subjects as if one group. One does not necessarily report in the write-up every statistic or graph one calculates if it does not prompt anything interesting to say about it, but even statisticians often comment that researchers often 'don't look at their data enough, but just want to do a significance test and get on to the next thing'.

DECISIONS ABOUT THE RIGHT GRAPHS AND DESCRIPTIVE STATISTICS

If you are not into one variable inferential stats (which we are omitting here), the choices to be made are simple:

Scale type of Variable:

	Interval	**Rank order**	**Categories of any sort**	**Counts in a continuum**
Graphic Presentation:	*histogram* of *scores*, frequency polygon	ordered list of cases	*bar chart*, pie chart (of *frequencies* or *percent*)	single bar
Centrality statistic:	*mean*, median score, modal score	median rank	*modal category*	frequency
Variation statistic:	*standard deviation*, range	quartile deviation	index of commonality	

-- Only the italicised ones are commonly met and will be dealt with here.
-- The modal category is simply the one with the most cases in it - the most popular one.
-- The mean (denoted by M or X-bar) is what we usually call the average in everyday English.
-- The standard deviation (SD) is a measure of the spread of scores. Roughly it is the average of the differences between each score and the mean score (see any stats book for the formula). So if everyone in a group scores the same, which will be the mean for the group, then the average of the differences of each score from the mean is 0 (SD=0; no variation). The more each score differs from the mean, the higher the SD gets, indicating more variation or 'disagreement' in the group. Usually one 'wants' small SDs.
-- Similar concepts to SD, calculated in various ways, are called 'error' and 'variance' in statistics.

Joke from WWW: Most of us have A Greater Than Average Number of Legs

The great majority of people have more than the average number of legs. Amongst the 57 million people in Britain there are probably 5,000 people who have only one leg. Therefore the average number of legs is

((5000 x 1) + (56,995,000 x 2)) / 57,000,000 = 1.9999123.

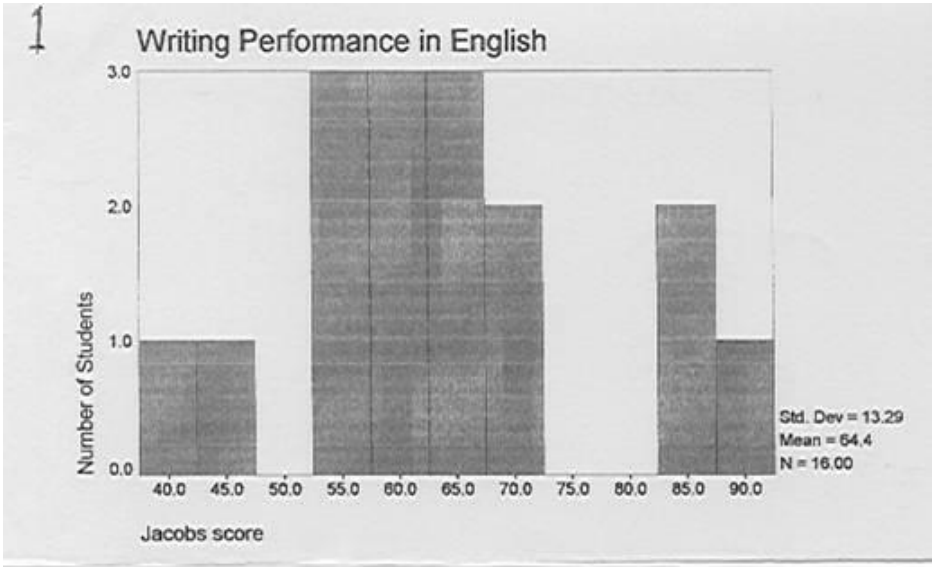
Since most people have two legs... need I say more?

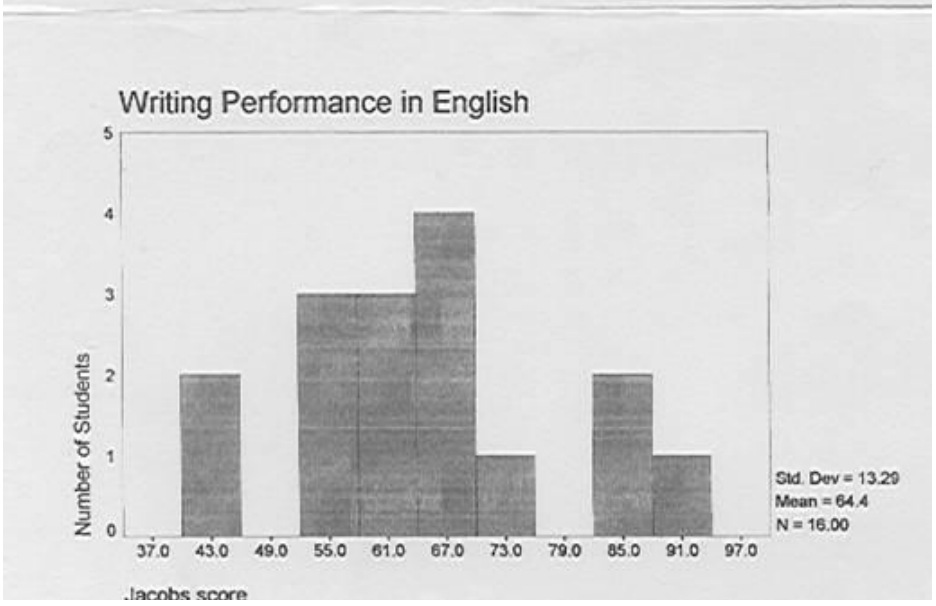
A FEW EXAMPLES OF SIMPLE GRAPHS: HOW TO MISLEAD

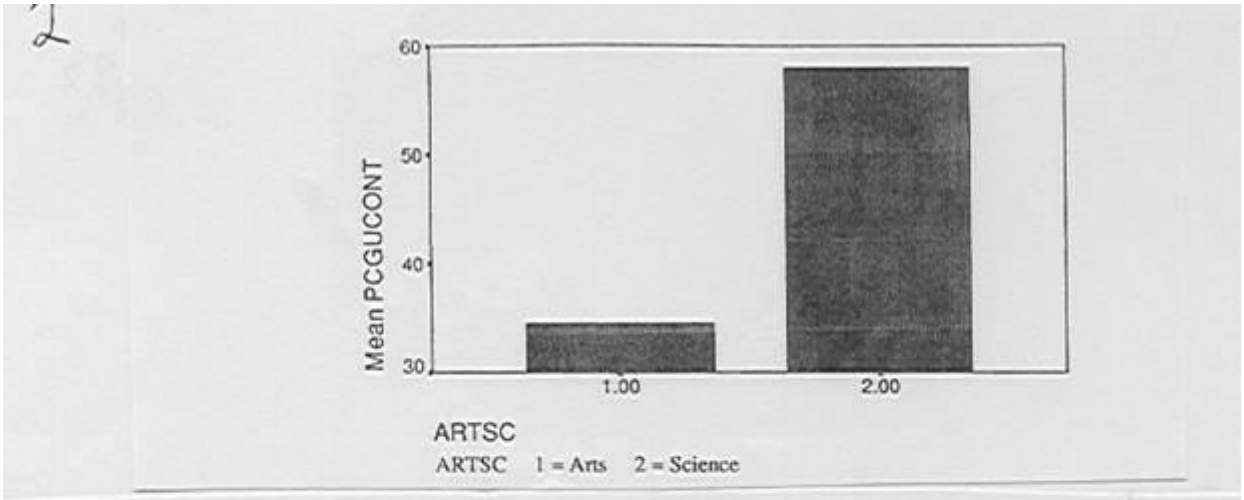
1. Shows two versions of a **histogram** of results for one group of 16 learners on one variable ('interval' scores for quality of each subject's written composition). Which version is better and why, or is neither optimal? What distinguishes a histogram from a bar graph/chart (seen in 2)? When to use each?

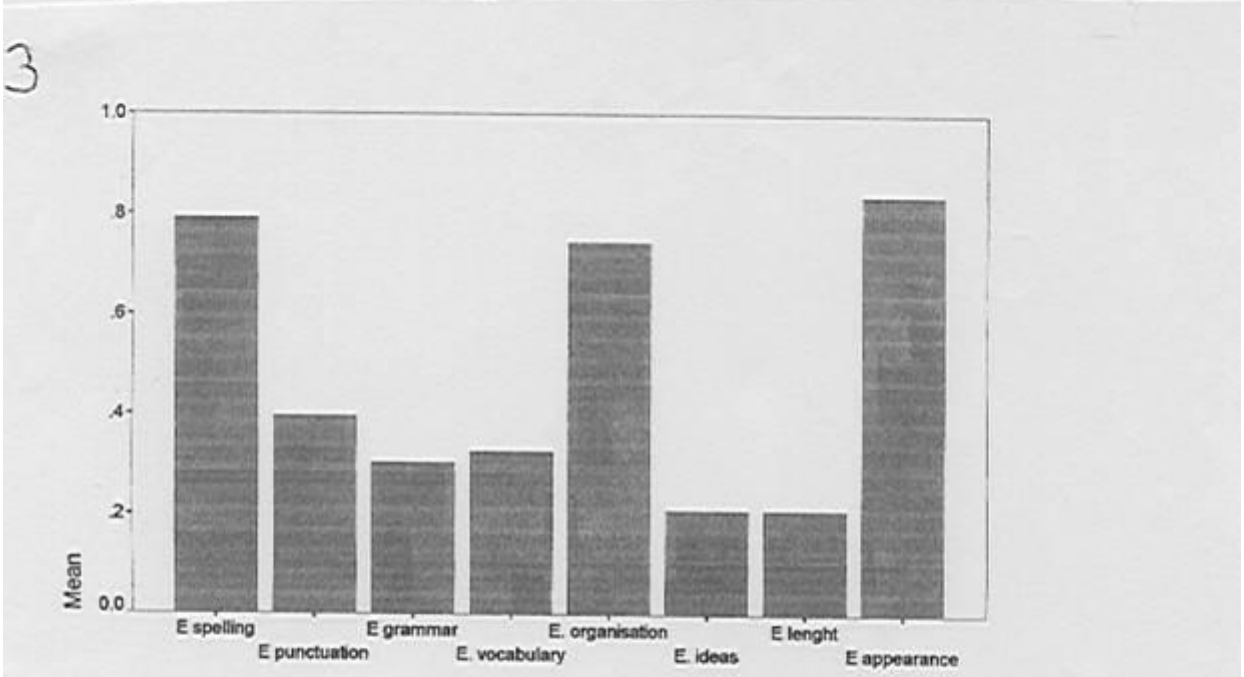
2. Is a **bar graph** (=bar chart) showing the broad subject specialism of participants in a study. I.e. it displays how all cases are categorised on a two category variable. How would you improve it for inclusion in a write-up?

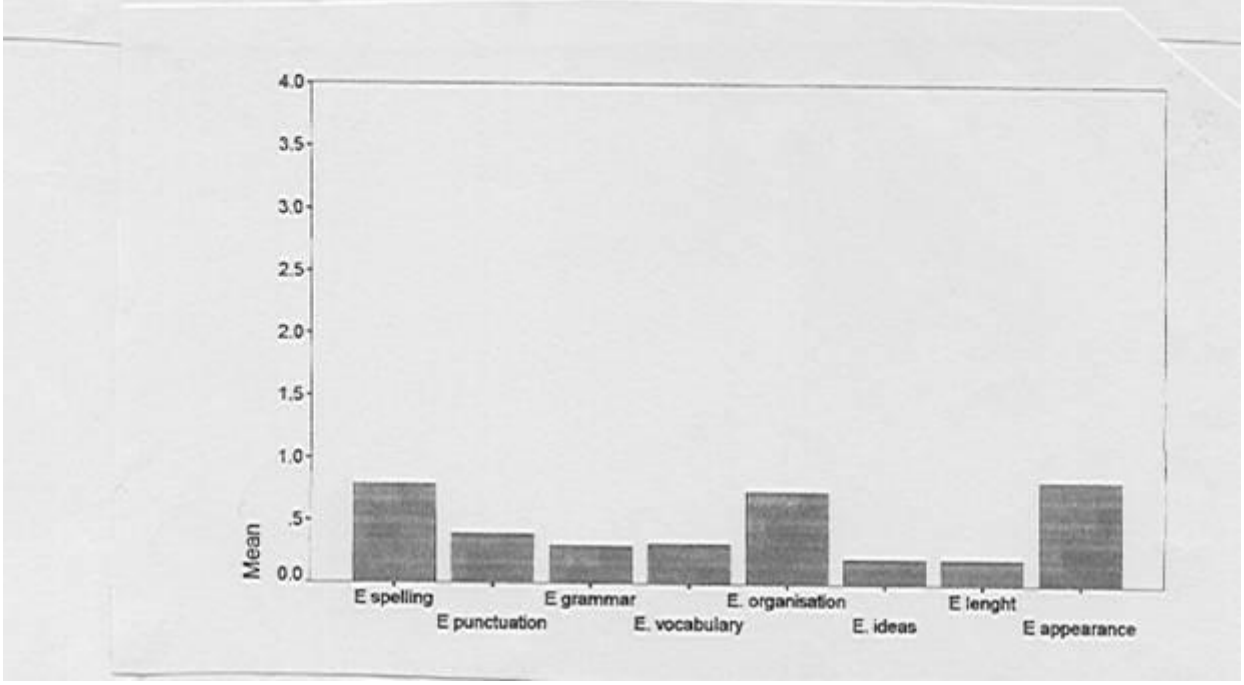
3. Shows two bar graphs for the same data – a set of several mean scores. One group of learners has given their ratings (on a five point scale) of how much they think eight different aspects of their compositions improved when done by word processing. The average ratings for each of these 8 variables are displayed together. Which is the better version and why?






60
50
40
30
Mean PCGUCONT
1.00
2.00
ARTSC
ARTSC 1 = Arts 2 = Science


1.0
.8
.6
.4
.2
0.0
Mean
E spelling
E punctuation
E grammar
E. vocabulary
E. organisation
E. ideas
E lenght
E appearance


4.0
3.5
3.0
2.5
2.0
1.5
1.0
.5
0.0
Mean
E spelling
E punctuation
E grammar
E. vocabulary
E. organisation
E. ideas
E lenght
E appearance

SIMPLE PERCENTAGES: HOW TO MISLEAD

1) Which sounds more impressive, A or B?

A) 2 out of 4 subjects agreed B) 50% of subjects agreed

A) 40 out of 80 subjects agreed B) 50% of subjects agreed

OK, but which result would you actually trust more? How should one report such results?

2) What is unclear? How to restate this better?

> In our survey we polled 50 people, though 10 declined to participate. …. 60% said yes to the question 'Do you like the English class?'…

3) Percentage scores versus group/aggregate percent.

Two ways of handling data arising from different numbers of potential occurrences for different people. Imaginary example of data where three subjects have been recorded in quasi-natural conversation, and counts have been made of their NS-like/correct use of third person *–s*.
Why do the percent differ in A and B? Which would the statistician prefer and why?

A) Analysis with subjects as cases: percentage scores and their mean

Case	**Correct**	**Incorrect**	**Total occurrences**	**Percent correct score**	**Mean percent correct**
Learner 1	12	12	24	50	
Learner 2	8	12	20	40	
Learner 3	3	9	12	25	
Total	23	33	56		

B) Analysis with occurrences as cases: group percent

		Total frequency	**Percent**
	Correct	23	41.1%
	Incorrect	33	58.9%
	Total occurrences	56	

PJS rev 05

 # The End .. with all wishes ☺