Appendix: A Mixed Methods Study

Classroom Culture

Carol Griffiths

The small-scale action research study on classroom culture which has been progressively used to illustrate the various steps throughout this book is an example of a mixed methods study in that it uses a questionnaire to gather quantitative data about student perceptions of impoliteness in a university classroom and also a grounded approach to qualitatively analyse the comments. Since the piecemeal presentation to which it has been subjected chapter by chapter may well have destroyed any sense of how it "hangs together", the full version is presented here for those who may be interested. Note that, since part of its intended purpose was to act as a model for how the students might go about their own research, there is no abstract, keywords, or introduction, and only a very brief literature review, all of which would need to be added or extended if it were intended for submission to a journal (which was never the case with this study).

Literature Review

Culture is always a difficult concept to define and is particularly hard to separate from concepts such as ethnicity and nationality (Zhou and Griffiths 2011). In its anthropological sense, culture is regarded as any of the customs, worldview, language, kinship systems, social organization, and other taken-for-granted day-to-day practices of a people which set them apart as a distinctive group (Scollon and Scollon 2000), whereas ethnicity refers to race (e.g. Maori) and nationality is essentially a political concept (what goes on one's passport).

A classroom is a specific sociocultural environment which has its own set of rules, language, expected behaviour, and standards of politeness (e.g. Griffiths et al. 2011). If these standards are not met, it may stop us from working well with others whose norms regarding what is polite or appropriate are different from our own (e.g. Neuliep 2003). It is therefore important to be clear about the behaviour expected, especially since these behaviours might not be identical across all contexts.

Research Question

This study asked the following questions:

What are the expected standards of behaviour in a Turkish university classroom?

Do these standards vary according to gender or nationality?

Research Context

Participants

There were 35 students who participated in this study. They were in the fourth year of a four-year degree in English Language Teaching (ELT), so they had already been at the university for some time, and might, therefore, have been expected to have clear ideas about the standard of behaviour expected. There were 13 male students and 22 female. Turkish students were a majority (N = 24), and there were 11 international stu-

dents (from Nigeria, Turkmenistan, Tajikistan, Iraq, Kirgizstan, Greece, and Kazakhstan).

Setting

The students were attending a Sociolinguistics course in the ELT Department in the Education Faculty at a private university in Istanbul, Turkey. Since the study was carried out in an actual classroom, where students were mainly concerned with passing their course (especially since many hoped this would be their last semester before graduation), the study had a multiple purpose: in addition to the research purpose, the topic was used to stimulate discussion, and also to act as a model for research projects in which students were engaged as part of their course.

Data Collection

As a first step, the issue of politeness was discussed in class, and students were asked to provide ideas about their perceptions of polite and impolite behaviour. They wrote these ideas on pieces of paper which were handed in anonymously. Based on this input, a questionnaire was then constructed (see Appendix for the questionnaire). A five-point Likert scale format was used, ranging from 5 = very impolite to 1 = not at all impolite. In addition, in order to obtain qualitative data, a column was provided where students were asked to comment and explain their opinions. In the interests of preserving privacy if students chose to remain anonymous, providing their names was optional, and in order to ensure consensual participation, students were asked to sign an agreement to using the data for research or publication purposes at the bottom of the questionnaire form.

Data Analysis

The questionnaire ratings were entered into the Statistical Package for Social Sciences (SPSS) and, since Likert-scale data are nonparametric, the data were analysed for median ratings. In addition, Mann–Whitney

U tests of difference for nonparametric data were used to investigate any differences in the ratings according to gender or nationality. The comments were also examined for ınsights which helped to explain the ratings.

Results

The median ratings results are set out in Table 1. According to these results, Items 7 and 8 are rated 5 ("very impolite"), whereas none of the items are rated 2 or 1 (in the "not impolite" range)

According to a Mann–Whitney U nonparametric test of difference, there were no statistically significant differences according to gender. According to nationality, only Item 1 showed a significant difference (p = .029), with international students considering coming late to class significantly more impolite than Turkish students.

A selection of the qualitative results from the comments are as follows:

1. Coming late

   - Being on time is your responsibility.
   - Causes lack of concentration.
   - If it is temporary OK, but if they keep coming late it is rude.

Table 1 Median ratings of questionnaire items (the higher the rating, the more impolite the behaviour is considered to be)

Item	Behaviour	Rating
1	Coming late	3
2	Using a phone	4
3	Talking while someone else is talking	4
4	Not paying attention	3
5	Sleeping	3
6	Eating or drinking	3
7	Using inappropriate terms of address	5
8	Using bad language	5
9	Wearing inappropriate clothing	3

2. Using a phone

   - It is disrespectful.
   - It disturbs others.
   - OK for emergency, but it causes students to lose their concentration

3. Talking while someone else is talking

   - We should behave as we want to be treated.
   - Plain rude.
   - Everybody needs to respect one another.

4. Not paying attention

   - This is disrespectful.
   - It distracts others
   - Don't attend the class if you won't pay attention anyway.

5. Sleeping

   - There is always coffee.
   - It is disrespectful.
   - Bedrooms are for sleeping, not classes.

6. Eating or drinking

   - Sometimes we don't get time for lunch or breakfast.
   - OK unless loud or smells bad.
   - Maybe he/she has an illness.

7. Using inappropriate terms of address

   - This is very rude.
   - It may cause unfriendly situation or arguments.
   - It can be very embarrassing, but it depends on the culture.

8. Using bad language

- This is impolite not just in class but everywhere.
- Can cause anger, losing concentration, and so on.
- Unrespect for the teacher.

9. Wearing inappropriate clothing

- It may be distracting. This is a school.
- It is a personal choice.
- It is important to be comfortable.

10. Any other behaviour you consider impolite

- Not bringing the right material.
- Not sitting upright.
- Chewing gum.
- Doing other work/reading in class.
- Being noisy.
- Leaving the class without permission.
- Attending to personal grooming (e.g. applying make-up, combing hair)

Discussion

It is interesting that all of the items in the scale were rated on the medium to very impolite side of the scale, with using inappropriate terms of address and using bad language being considered most impolite (median = 5). Using a phone and talking while someone else is talking are also considered impolite (median = 4). Some participants felt that, given Istanbul's notorious traffic, coming late was sometimes unavoidable, while eating or drinking in class might be all right under some circumstances as long as it did not inconvenience others, since there are days when students start early and finish late with no breaks between classes. A reasonable degree of tolerance seemed to be extended to standards of dress, as long

as it was not "embarrassing" or "distracting", and the question was also raised by some students that it was not easy to achieve a definition of what was "appropriate", a point that was also made regarding terms of address (these points were discussed in class as part of a follow-up to the study). Several other impolite behaviours were suggested, including failure to bring the correct material, chewing gum, and attending to personal grooming.

Another interesting feature of the results is the unanimity across learner variables. We might have expected males and females to have different opinions about some of these behaviours, but this did not prove to be the case. Furthermore, although there was one statistically significant difference according to nationality, suggesting that Turkish students are less concerned about punctuality than their international classmates, the remainder of the differences according to nationality were not significant.

Implications

Once the results had been found and the study written up, it was distributed to the class for discussion. There was some amusement with some of the findings (e.g. about eating, drinking, or sleeping in class), and some embarrassment from some who recognized impolite behaviours in themselves (e.g. consistently coming late and entering in such a way as to distract classmates, using phones during class). In the weeks following the study, there was some evidence that students made some effort to control their own negative behaviour (e.g. by being more punctual, turning off phones when they arrived in the room) and some also seemed to be more inclined to attempt to control classmates for impolite behaviour (e.g. if they were talking while someone else—either the teacher or one of their own classmates during a presentation). In other words, this study seemed to have some effect with raising students' own awareness of the effect of their behaviour on others, and to result in more willingness to engage in polite behaviour, which, in turn, seemed to have a positive effect on classroom dynamics.

Suggestions for Further Research

There are a number of ways in which this study might be extended:

- An interesting direction would be to give the survey to teachers and compare their perceptions with those of students.
- Also, interviews could be conducted for a more in-depth qualitative perspective.
- Some of the other behaviours the students considered impolite might be added to the questionnaire.

Conclusion

Overall, there is agreement that four out of nine (44.4%) of the behaviours listed in the questionnaire are impolite or very impolite. Students did not express strong opinions about the remaining five items (55.6%), but none of the items was rated polite. Perhaps if students are made aware of these findings it might help to promote a more harmonious classroom atmosphere which is likely to promote more effective learning.

References

Griffiths, C., Raman, T., & Mannes, M. (2011). Have I offended you? Appropriate nonverbal communication in the international classroom. *Journal of International Education and Business, 3*, 7–29.

Neuliep, J. (2003). *Intercultural communication: A contextual approach.* New York: Houghton Mifflin.

Scollon, R., & Scollon, S. (2000). *Intercultural communication: A discourse approach.* Beijing: Foreign Language Teaching and Research Press.

Zhou, C., & Griffiths, C. (2011). Intercultural communicative competence. *English Language and Literature Studies, 1/2*, 113–122

Appendix: Classroom Culture Questionnaire

Gender: M F Nationality:

Please mark the following behaviour according to how impolite you think it is. Please also add a comment explaining your opinion.

5 = very impolite 4 = impolite 3 = no strong opinion 2 = not impolite 1 = not at all impolite

Item	Behaviour	Rating	Comment: please explain your rating
1	Coming late		
2	Using a phone		
3	Talking while someone else is talking		
4	Not paying attention		
5	Sleeping		
6	Eating or drinking		
7	Using inappropriate terms of address		
8	Using bad language		
9	Wearing inappropriate clothing		
	Any other behaviour you consider impolite		

I agree to the use of this data for research or publication purposes:

Glossary

Abstract A short summary of the main ideas and findings placed at the beginning of an article.

Action research First used by Lewin (1946), the term action research involves an active investigation of a problem, puzzle, question, or issue by a researcher or researchers (often the teacher/s) who is/are a participant or participants in the given context (Burns, 2010; Lewin, 1946).

ANOVA (analysis of variance) A statistical procedure for testing whether the difference between parametric (q.v.) variables is significant (q.v.)

APA This refers to a style established by the American Psychological Association, commonly used in social science publications.

Case study This kind of research takes an in-depth look at a specific definable unit, for example, an individual, a particular class, or a specific school.

Categorical data Another term for nominal data (q.v.) where the data are divided into categories.

Collaborative action research Teachers working collaboratively to do action research with the purpose of investigating and reflecting critically on their teaching practices.

Confidence interval The confidence interval indicates a certain degree of probability that we can have confidence that a given result is more than mere chance. For instance:

 $p<.05$ indicates that we can be 95% confident that the result is significantly more than would be found just by chance. Even better, $p<.01$ indicates a 99% level of confidence that the result is significantly more than would be found just by chance.

Context The context refers to the environment (e.g. a nation, a district, a school, a workplace, or a classroom) in which a given phenomenon occurs. Context may include both the setting where the phenomenon occurs and the people involved.

Continuous data Operating over a range, for example, test scores from 0% to 100%.

Convenience sampling This refers to a group of participants chosen for a study based on their convenient availability. For instance, if I want to investigate the effect of strategy training, I might decide to study my own class on the basis that they are easily available, whereas other groups might not be so convenient.

Correlation This refers to a statistical procedure to measure the strength of the relationship between variables. Correlations can vary between 1 (meaning there is a perfect correlation, typically found only between items that actually are the same), 0 (which means there is no correlation), and −1 (meaning that the correlation is negative, so that as one variable increases the other decreases). Two well-known tests of correlation are:

- **Pearson product–moment correlation** or Pearson's **r** for numerical (q.v.), normally distributed (q.v.) data such as test scores etc.
- **Spearman's rank–order correlation** or **Spearman's rho** (r_s) for data which is not normally distributed (q.v.) or ordinal (q.v.) data such as obtained from Likert-type questionnaires (q.v).

Criticality The practice of 'socially situated reflection and evaluation' (Banegas, 2016, p. 1) which involves creating multiplicity in perspectives while elaborating on issues, including self-criticism.

Cross-sectional These kinds of studies look at the relationships between/among various factors at a particular point in time (cf. longitudinal).

Deductive analysis Pre-setting the potential themes that are to be deduced from the data.

Descriptive statistics This term is used to include procedures such as mean, median, mode, frequencies, and percentages which are used to describe the results of a particular study.

Differences Whereas tests of correlation test how variables relate to each other, tests of difference examine whether given variables (e.g. reported strategy use) vary according to different groups (e.g. male/female). Common tests of difference include:

- **Student's *t*-test**—named after William Gosset who introduced the procedure under the pseudonym "Student", this test is used for numerical (q.v.), normally distributed (q.v.) data such as finding the differences in test scores according to gender.
- **Mann–Whitney U**—this test is used for two independent groups (e.g. male/female) when the dependent variable is either ordinal (q.v.), for example, reported strategy use, or numerical (q.v.) but not normally distributed (q.v.).
- **Kruskal–Wallis**—this test is used for groups (e.g. several different nationalities) when the dependent variable is either ordinal (q.v.), for example, reported strategy use, or numerical (q.v.) but not normally distributed (q.v.). This test is often used for nonparametric (q.v.) data where an ANOVA (q.v.) would be used for parametric (q.v.) data.

Documentation There are numerous documents which might be kept to serve as evidence for research studies, including:

- lesson plans
- teaching materials
- student work
- journals/diaries/blogs
- pictures
- tests and results

Effect size This refers to the strength of one variable's effect on another, for example, what is the effect of motivation on target language proficiency development.

Ethics Ethical behaviour in research refers to a number of issues, including:

- **Informed consent:** participants have a right to be informed about the purpose of the research and how the data they provide will be used.
- **Voluntary participation:** participants have a right to understand that they do not have to participate, and that they can withdraw at any time if they wish to do so.

- **Anonymity:** participants do not have to identify themselves if they do not wish to do so.
- **Confidentiality:** information given during the data collection process must remain confidential.
- **Storage of data:** data obtained during the collection process must be stored securely and kept for an approved period of time.
- **Absence of threat:** participants must be assured that the data they provide will have no effect on their results or cause any other disadvantage.
- **Professional distance:** teachers should take care to maintain the teacher–student relationship at all times. (**See Chap. 5 for more details**.)

Ethnography Ethnographic studies examine a particular phenomenon which occurs in a specific cultural context.

Experimental A true experimental study involves using:

- random assignment
- control versus experimental groups
- application of some kind of intervention
- use of pre- and post-tests to examine the effect of the intervention.

Exploratory action research A form of research requiring extensive initial discussion of practice before attempting to make further practical and pedagogical changes.

Exploratory practice A form of research requiring exploring the context with co-researchers such as learners through normal pedagogical practices as research tools.

Factor analysis A factor analysis is used to determine whether particular items in a data set "hang together". There are two types of factor analysis:

- **An exploratory factor analysis** uses the factor analysis procedure to identify factor groupings, for example, in a newly constructed instrument.
- **A confirmatory factor analysis** is used to determine whether factors do, in fact, hang together according to some pre-determined grouping, for example, an existing questionnaire.

Feasibility This refers to how practical a given research idea may be in a given context.

Grounded theory Grounded theory describes a research approach which does not begin with hypotheses, but where the conclusions emerge "from the ground up" as a result of the investigation. Typically, there are three stages of coding involved with grounded theory:

- **Open coding:** At this stage, the data is examined for salient themes.
- **Axial coding:** At this stage, the themes identified in Stage 1 (which can be numerous) are grouped around central axes.
- **Selective coding:** At this final stage, a single overarching theme is identified.

Hawthorne effect Also called observer paradox (Labov, 1972), the Hawthorne effect (after the place where it was first documented, Landsberger, 1958) explains the phenomenon that, although an observer may go into a situation in order to observe natural behaviour, the very fact of being there is likely to change normal patterns of behaviour.

Hypothesis/es A hypothesis states the researcher's prediction/s concerning the finding/s that a given study is likely to arrive at. **A null hypothesis (q.v)** proposes that there will be no difference between the groups being researched. If a difference is found, the null hypothesis will be rejected and the alternative hypothesis will be supported.

Inductive analysis This involves inducing conclusions from the data rather than deducing conclusions according to pre-set criteria.

Inferential statistics This term is used to include procedures such as correlation, differences, and effect size which are used to infer generalizations beyond the immediate sample.

Instrumentation Instrumentation refers to the materials or instruments (e.g. a questionnaire, interview schedule, observation checklist) used in the course of the study.

Inter-rater reliability This is important for qualitative studies to help reduce the risks of subjectivity if just one researcher's assessments are recorded. Inter-rater reliability is often expressed as a percentage, which represents the total number of the items on which the different raters agree divided by the total number of items identified. Any disagreements need to be resolved by negotiation.

Interviews The interview technique involves canvassing the opinions of an individual or group of individuals. Interviews can be:

- **Structured** (where the questions are decided beforehand and only those questions are used)
- **Unstructured** (where the interview proceeds according to issues which arise spontaneously)
- **Semi-structured** (where a set of pre-planned questions forms the basis of the interview, but there is some flexibility to discuss other issues or insights as they arise)

Interval data Where values have regular intervals between them, for example, children's ages in years, worker's income in dollars.

Introspection Introspective methods involve asking participants to reflect on their thinking processes and to report on these while they are performing a task. Often also called **think-aloud procedures** (q.v.).

Keywords The purpose of keywords is so that key topics will appear on a database for someone looking for this subject. As such, they are an important aid to citability.

Kolmogorov–Smirnov test This test checks whether data are normally distributed around a mean (see also Shapiro–Wilk test). If the significance value for an item is less than .05, distribution is not normal. This is important, since parametric tests such as Pearson product–moment correlation, t-tests, and ANOVAs (q.v.) assume normal distribution, and, if this is not the case, nonparametric equivalents such as Spearman rank order correlation, Mann–Whitney U or Kruskal–Wallis H (q.v.) should be used. The Kolmogorov–Smirnov is considered more reliable for smaller numbers (less than about 100) than the Shapiro–Wilk.

Kruskal–Wallis H test of difference This is a nonparametric (q.v.) test useful for testing whether any differences among several data sets which are non-numerical and/or not normally distributed are significant. It is often reckoned to be the nonparametric equivalent of an ANOVA (q.v.).

Lesson study A form of reflection supported by peer observation to develop better understanding of actual classroom settings at the moment of teaching as well as student learning.

Likert-type questionnaire. This kind of instrument, invented by psychologist Rensis Likert (1932), involves asking participants to rate questionnaire items on a continuum, typically, for instance, from strongly agree to strongly disagree, from always to never or some similar range which expresses opinion or attitude.

Literature review A summary of the existing published literature on a given topic.

Longitudinal Taking place over a period of time (cf. cross-sectional).

Mann–Whitney U This is a nonparametric (q.v.) test useful for testing whether there are any significant differences between two data sets which are non-numerical and/or not normally distributed. It is often reckoned to be the nonparametric equivalent of a t-test (q.v.).

MANOVA (Multivariate analysis of variance). A statistical procedure for testing whether the difference among multiple parametric variables is significant (cf. ANOVA, Kruskal–Wallis H).

Mean The mean (also called "average") is obtained by adding all numbers together and dividing by the number of the numbers. So if a group of students score 10, 8, 8, 7, 5, 4, 2 on a test, the mean is the total (=44) divided by how many numbers (=7), so the mean=6.28.

Median The median is the number at the midpoint of a set of numbers. In the example above for the mean, the median is 7.

Member-checking Asking for the feedback of the people who provided data for the analysed data in order to increase the reliability and validity.

Methodology The methodology refers to how a researcher went about conducting the study, and includes details about the participants, the context, how the data were collected and analysed, the instruments that were used, ethical procedures, and so on.

Mixed methods A mixed method study is one which, rather than keeping to just one method (e.g. a questionnaire or interviews), uses more than one method (e.g. a questionnaire, followed up by interviews). The advantage of such methods is that they provide triangulation (q.v.).

Mode The mode of a set of numbers is the one which occurs most frequently. In the case of the numbers exemplifying "mean" above, the mode is 8.

Narrative inquiry This kind of methodology gathers data by means of asking participants to narrate their experiences, feelings, opinions, and so on. These data are then analysed by means of qualitative methods such as grounded theory (q.v.).

Nominal/categorical data This occurs where variables are given numbers for the sake of convenience for entering them into a computer programme. For instance, males=1, females=2. Or Chinese=1, Indians=2, Turkish=3, and so on. These values are purely arbitrary, and have no actual numerical value. As a result, they cannot be analysed by either parametric (e.g. Pearson's product–moment correlation) or nonparametric (e.g. Spearman's rho) tests, though they can be used as grouping variables in tests of difference.

Nonparametric Nonparametric data are not normally distributed (q.v.) and often non-numerical. A common example of these types of data include Likert-type questionnaires (q.v.), where the numbers actually represent strength of opinions (e.g. "strongly disagree") rather than real numbers.

Normal distribution If numbers are normally distributed, they form a symmetrical bell-shaped curve when graphed. Normality of distribution can be tested by means of a Kolmogorov–Smirnov or a Shapiro–Wilk (q.v.) on SPSS.

Null hypothesis (H_0). This is a statement that there is no correlation or difference between given variables. If a correlation or difference is found, the null hypothesis is rejected, and the alternative hypothesis is accepted.

Numerical data Are real numbers (e.g. ages, incomes, test scores), as opposed to ordinal (q.v.) or nominal (q.v.) data.

Observation One of the ways of understanding practices as well as learning about learners and teachers through direct access to the classroom. It provides first-hand data to the researchers or other teachers.

Observer paradox See Hawthorne effect.

Ordinal data This kind of data is generated from procedures such as Likert-type questionnaires. The numbers assigned to judgements such as "strongly disagree" are clearly not actual numbers: they merely indicate strength of opinion about a given question/statement. Furthermore, we cannot assume that someone who assigns a "4" to a given item is twice as much in agreement as someone who only gives it a "2": this simply makes no sense. In other words, this kind of data is non-numerical, and should be analysed using non-parametric tests such as Spearman's rho (for correlations, q.v.) and Mann–Whitney U or Kruskal–Wallis (for differences, q.v.).

Parametric This kind of data consists of real numbers, which operate within set parameters. For instance, we can say that a child of 4 is half as old as a child of 8. Someone who earns \$50,000 per year earns twice as much as someone who earns \$25,000. A student who gets 60% on a test gets 2/3 of what a student who gets 90% earns. As long as this kind of data is normally distributed (q.v,) it can be analysed using parametric tests such as Pearson's product–moment correlation, Student's *t*-test of difference, or ANOVA (q.v.).

Participants Participants are those who take part in a study, perhaps by answering questionnaires, attending interviews, and so on.

Pearson product–moment correlation A statistical test for calculating the relationship between two numerical variables (e.g. exam scores, age).

Phenomenology This is the study of a specific phenomenon (e.g. unmotivation, age or gender differences, autonomy, metacognition.) in a given context.

Piloting The term used for trialing an instrument or research procedure in order to identify any problems before using it with the main target participants.

Population A group of people who share some particular characteristic (e.g. university students, young readers, older language learners, female language learners) from which a sample (q.v.) is drawn.

Potentially exploitable pedagogic activities (PEPAs) Classroom activities that integrate teaching practices and research purposes especially in collecting data. These activities benefit students rather than use them only for eliciting research data.

Probability In statistics, this term refers to the probability that a given effect has occurred by more than just chance. In social sciences, 95% probability ($p < .05$) is generally accepted as the lower threshold to conclude that a result is statistically significant.

Problematize Investigating an issue in the classroom which is not necessarily deficient but may be developed or understood from different perspectives.

Purposive sampling Purposive sampling occurs when participants are selected for some particular reason, such as their gender, age, achievement level or nationality (cf. convenience and random sampling).

Qualitative research This refers to studies which result in data which cannot be expressed in numerical terms, such as interviews, observations, case studies, think-alouds, and so on.

Quantitative research This refers to studies which produce data which can be expressed as numbers, such as Likert-type questionnaires (q.v.).

Quasi-experimental This refers to the kind of study which attempts to assess the effects of a particular treatment, but where one or more of the conditions for a true experiment (e.g. random assignment, q.v.) are not met.

Questionnaire An instrument for collecting data, with open and/or closed items. Questionnaires can be paper-and-pencil, or, increasingly commonly, they can be distributed electronically.

Random sampling This occurs where participants are selected for a study purely by chance. (cf. convenience and purposive sampling).

Reflective practice Reflecting upon classroom practices in order to evaluate their efficiency in a systematic way.

Regression A regression analysis is a statistical technique for using independent variables to predict a dependent variable. For instance, if a student scores 30% on one assessment, and 40% on another, a regression analysis could be used to predict the likelihood of him or her obtaining a passing grade (60%). In an example such as the above where there is more than one independent variable, it is called multiple regression.

Reliability This refers to the consistency of the data. There are two main types: internal and external (see Chap. 1: Quality of the data).

Replication Replication occurs when a study which has been carried out in one context is repeated in a different environment in order to see if similar results are produced.

Reports Reports are constructed by means of gathering and analysing available documents, such as journals, blogs, class records, and so on.

Sample A sample is a section of a larger population (q.v.) which is used (usually for practical reasons, since surveying an entire population is often difficult if not impossible) to gather information about a particular phenomenon. For instance, if we want to find out about first language maintenance among immigrant children, we might begin by surveying children at just one school, or even just one class, and begin to draw conclusions from there.

Setting This term refers to the place or location where particular phenomena occur.

Shapiro–Wilk test Like the Kolmogorov–Smirnov test (q.v.) this test checks whether data are normally distributed around a mean. If the significance value for an item is less than .05, distribution is not normal. This is important, since parametric tests such as Pearson product–moment correlation, t-tests, and ANOVAs (q.v.) assume normal distribution, and, if this is not the case, nonparametric equivalents such as Spearman rank order correlation, Mann–Whitney U or Kruskal–Wallis (q.v.) should be used. The Shapiro–Wilk test is usually considered to be more reliable than the Kolmogorov–Smirnov for larger numbers (100+).

Significance The significance (sig.) figure tells us the probability that a particular result is likely to be more than chance. In social sciences, the lower threshold is generally accepted as 95% (p.<05).

Spearman rank order correlation This is a statistical test for calculating the relationship between two nonparametric (q.v.) variables (e.g. expressions of strength of agreement on a Likert-type questionnaire).

Standard deviation (SD) A measure of the degree to which scores deviate from the mean. If the SD is low, it indicates scores are close together. If it is large, it means scores vary considerably.

Sustainability Sustainability refers to creating potential longer-term impacts of the benefits of action research on teachers.

Teacher autonomy Teachers' ability to make decisions on their own professional development and create independent and interdependent opportunities for themselves.

Teacher research A form of research carried out by teachers in their own working context in order to improve their practices and develop their understanding.

Think-aloud The think-aloud technique is a form of introspection (q.v.) where participants are asked to verbalize their thinking processes for later analysis.

Triangulation Where different methods of data collection and analysis are used to cross-check the findings.

t-test This is a parametric (q.v.) statistical procedure for determining whether any differences between numerical, normally distributed variables are significant (q.v.).

Validity Validity refers to the extent to which a procedure measures what it is supposed to be measuring. There are several different types of validity, including construct/concept, content, convergent, concurrent, criterion, external, face, internal and predictive (see Chap. 1: Quality of the data).

References

Allwright, D. (1997). Quality and sustainability in teacher-research. *TESOL Quarterly, 31*(2), 368–370.

Allwright, D. (2000). Exploratory practice: An "appropriate methodology" for language teacher development. In *8th IALS symposium for language teacher educators,* Edinburgh, Scotland.

Allwright, D. (2003). Exploratory practice: Rethinking practitioner research in language teaching. *Language Teaching Research, 7*(2), 113–141.

Allwright, D. (2005). Developing principles for practitioner research: The case of exploratory practice. *The Modern Language Journal, 89*(3), 353–366.

Allwright, D., & Hanks, J. (2009). *The developing language learner*. Hampshire: Palgrave Macmillan UK.

Allwright, D., & Lenzuen, R. (1997). Exploratory practice: Work at the cultural inglesa, Rio de Janeiro, Brazil. *Language Teaching Research, 1*(1), 73–79.

Ary, D., Jacobs, L. C., Sorensen, C., & Walker, D. (2010). *Introduction to research in education.* Wadsworth: Cengage Learning.

Bailey, K. (2001). Action research, teacher research, and classroom research in language learning. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (3rd ed., pp. 489–498). Boston: Heinle & Heinle.

Bailey, K., Curtis, A., & Nunan, D. (2001). *Pursuing professional development.* Boston: Heinle & Heinle.

Banegas, D. L., & de Castro, L. S. V. (2016). Criticality. *ELT Journal*, ccw048.

Barkhuizen, G. (2009). Topics, aims and constructs in English teacher research: A Chinese case study. *TESOL Quarterly, 43*(1), 113–125.

Benson, P. (2000). Autonomy as a learners' and teachers' right. *Learner Autonomy, Teacher Autonomy: Future Directions, 3*(2), 111–117.

Besser, S., & Chik, A. (2014). Narratives of second language identity amongst young English learners in Hong Kong. *ELT Journal, 68*(3), 299–309.

Borg, S. (2010). Doing good quality research. *JACET Journal, 50*, 9–13.

Borg, S. (2011). Doing action research in English language teaching. A guide for practitioners. *ELT Journal, 65*(4), 485–487.

Borg, S. (2013). *Teacher research in language teaching: A critical analysis*. Cambridge: Cambridge University Press.

Borg, S. (Ed.). (2015). *Professional development for English language teachers: Perspectives from higher education in Turkey*. Ankara: British Council.

Bright, B. (1996). Reflecting on 'reflective practice'. *Studies in the Education of Adults, 28*(2), 162–184.

Burns, A. (1999). *Collaborative action research for English language teachers*. Cambridge: CUP.

Burns, A. (2009). Action research in second language teacher education. In *The Cambridge guide to second language teacher education* (289–297). New York: Cambridge University Press.

Burns, A. (2010). *Doing action research in English language teaching: A guide for practitioners*. New York: Routledge.

Burns, A. (2015). Renewing classroom practices through collaborative action research. In K. Dikilitaş, R. Smith, & W. Trotman (Eds.), *Teacher-researchers in sction* (pp. 9–17). IATEFL: Faversham.

Burns, A., & Hood, S. (1997). *Teachers' voices 2: Teaching disparate learner groups*.

Burton, J., & Mickan, P. (1993). Teachers' classroom research. In J. Edge & K. Richards (Eds.), *Teachers develop teachers research* (pp. 113–121). Oxford: Heinemann.

Cajkler, W., Wood, P., Norton, J., & Pedder, D. (2013). Lesson study: Towards a collaborative approach to learning in initial teacher education? *Cambridge Journal of Education, 43*(4), 537–554.

Carr, W., & Kemmis, S. (2003). *Becoming critical: Education knowledge and action research*. New York: Routledge.

Çelik, Ç., & Dikilitaş, K. (2015). Action research as a professional development strategy. In S. Borg & S. Sanchez (Eds.), *International perspectives on teacher research* (pp. 125–138). Palgrave: Basingstoke.

Child, A. J., & Merrill, S. J. (2003). Professional mentors' perceptions of the contribution of school/HEI partnerships to professional development and school improvement. *Journal of In-Service Education, 29*(2), 315–324.

Cochran-Smith, M., & Lytle, S. L. (1999). The teacher research movement: A decade later. *Educational Researcher, 28*(7), 15–25.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Lawrence Erlbaum Associates.

Creswell, J. W. (2014). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Essex: Pearson.

Crookes, G. (1993). Action research for second language teachers: Going beyond teacher research. *Applied Linguistics, 14*(2), 130–144.

Dikilitaş, K. (Ed.). (2012). *Teacher-research studies at foreign language school: Inquiries from teacher perspectives* (Vol. 1). Ankara: Nobel.

Dikilitaş, K. (Ed.). (2013). *Teacher-research studies at foreign language school: Inquiries from teacher perspectives* (Vol. 2). Ankara: Nobel.

Dikilitaş, K. (Ed.). (2014a). *Professional development through teacher research*. İzmir: Gediz University Press.

Dikilitaş, K. (2014b). Introduction. In K. Dikilitaş (Ed.), *Professional development through teacher research*. İzmir: Gediz University Press.

Dikilitaş, K. (2015). Teacher research for instructors. In S. Borg (Ed.), *Professional development for English language teachers: Perspectives from higher education in Turkey* (pp. 27–33). Ankara: British Council, Turkey.

Dikilitaş, K., & Mumford, S. E. (2016). Supporting the writing up of teacher research: Peer and mentor roles. *ELT Journal, 70*(4), 371–381.

Dikilitaş, K., & Wyatt, M. (2017). *Learning teacher-research-mentoring: Stories from Turkey*. Manuscript submitted for publication.

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.

Edge, J. (Ed.). (2001). *Case studies in TESOL practice: Action research*. Alexandria: TESOL.

Edwards, E., & Burns, A. (2016). Language teacher action research: Achieving sustainability. *ELT Journal, 70*(1), 6–15.

Farrell, T. S. (2008). Critical incidents in ELT initial teacher training. *ELT Journal, 62*(1), 3–10.

Freeman, D. (1996). Redefining the relationship between research and what teachers know. In K. M. Bailey & D. Nunan (Eds.), *Voices from the language classroom* (pp. 88–115). New York: Cambridge University Press.

Freeman, D. (1998). *Doing teacher research*. Boston: Heinle and Heinle.

Gibbs, G. (1988). *Learning by doing: A guide to teaching and learning methods.* Oxford: Further Education Unit/Oxford Polytechnic.

Gieve, S., & Miller, I. K. (2006). What do we mean by 'quality of classroom life'? In *Understanding the language classroom* (pp. 18–46). New York: Palgrave Macmillan UK.

Glaser, B., & Strauss, A. (1967). *The Discovery of grounded theory.* New York: Aldine.

Griffiths, C. (2003). *Language learning strategy use and proficiency: The relationship between patterns of reported language learning strategy (LLS) use by speakers of other languages and proficiency with implications for the teaching/learning situation.* Ph.D. thesis, University of Auckland, New Zealand. http://hdl.handle.net/2292/9

Griffiths, C. (2004). Language learning strategies: Theory and research. *Occasional Paper no.* 1. Auckland, New Zealand, Centre for Research in International Education. http://www.crie.org.nz

Griffiths, C. (Ed.). (2008). *Lessons from good language learners.* Cambridge: Cambridge University Press.

Griffiths, C. (2015a). *Dynamic strategy development and progress in language learning.* www.carolgriffiths.net

Griffiths, C. (2015b). What have we learnt from good language learners? *ELT Journal, 69*(4), 425–433.

Griffiths, C. (2016). Strategies for developing English language writing skills. *Asian EFL (online).*

Griffiths, C., & İnceçay, G. (2015). Styles and style-stretching: How are they related to successful learning? *Journal of Psycholinguistic Research, 45*, 599–613.

Griffiths, C., & Özgür, B. (2013). Second language motivation. *Procedia – Social and Behavioral Sciences, 70*, 1109–1114.

Günbay, E. B., & Aydemir, G. (2015). Tackling speaking challenges faced by low-level learners of English through consultation with students. In K. Dikilitaş, R. Smith, & W. Trotman (Eds.), *Teacher-researchers in action* (pp. 97–108). IATEFL: Faversham.

Güzel, M. (2015). Hey, teacher, leave those kids alone! Beliefs of students and instructors about learner autonomy. In K. Dikilitaş, R. Smith, & W. Trotman (Eds.), *Teacher-researchers in action* (pp. 189–206). IATEFL: Faversham.

Halai, A. (2006). Mentoring in-service teachers: Issues of role diversity. *Teaching and Teacher Education, 22*(6), 700–710.

Hanks, J. (2009). Inclusivity and collegiality in exploratory practice. In *Researching language teaching and learning: An integration of practice and theory* (pp. 33–55). New York: Peter Lang.

Hanks, J. (2014). 'Education is not just teaching': Learner thoughts on exploratory practice. *ELT Journal*, ccu063.

Hanks, J. (2015). Language teachers making sense of exploratory practice. *Language Teaching Research*. doi:10.1177/1362168814567805.

Hendricks, C. (2009). *Improving schools through action research: A comprehensive guide for educators* (2nd ed.). Boston: Allyn & Bacon.

Hobson, A. J., & Sharp, C. (2005). Head to head: A systematic review of the research evidence on mentoring new head teachers. *School Leadership & Management, 25*(1), 25–42.

Horwitz, E. (1987). Surveying student beliefs about language learning. In A. Wenden & J. Rubin (Eds.), *Learner strategies in language learning* (pp. 119–132). London: Prentice Hall.

Horwitz, E., Horwitz, M., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125–132.

Hu, H., & Nassaji, H. (2014). Lexical inferencing strategies: The case of successful versus less successful inferencers. *System, 45*, 27–38.

Kordia, S. (2015). From TEFL to ELF-aware pedagogy: Lessons learned from an action-research project in Greece. In K. Dikilitaş, R. Smith, & W. Trotman (Eds.), *Teacher-researchers in action* (pp. 235–262). IATEFL: Faversham.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania.

Landsberger, H. (1958). *Hawthorne revisited*. New York: Ithaca.

Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues, 2*(4), 34–46.

Lewis, K. A. (2009). Adult learners' perceptions of the incorporation of their L1 in foreign language teaching and learning. *Applied Linguistics, 30*(2), 216–235.

Lewis, C., Perry, R., Hurd, J., & O'Connell, M. P. (2006). Lesson study comes of age in North America. *Phi Delta Kappan, 88*(4), 273.

Lieberman, A. (1995). Practices that support teacher development: Transforming conceptions of professional learning. *Phi Delta Kappan, 76*, 591–596.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140*, 1–55.

Lin, M., Li, J., Hung, P. and Huang (2014). "Blogging a journal: Changing students' writing skills and perceptions". *ELT Journal*, 68/4, 422-431.

Little, D. (1995). Learning as dialogue: The dependence of learner autonomy on teacher autonomy. *System, 23*(2), 175–181.

Livingston, K., & Shiach, L. (2013). *Teaching Scotland's future*. Mentoring Pilot Partnership Project, Final report.

Loughran, J. J. (2002). Effective reflective practice in search of meaning in learning about teaching. *Journal of Teacher Education, 53*(1), 33–43.

Lynch, B. K. (1996). *Language program evaluation: Theory and practice*. Cambridge/New York: Cambridge University Press.

MacLean, M. S., & Mohr, M. M. (1999). *Teacher-researchers at work*. National Writing Project, University of California, 2105 Bancroft,# 1042, Berkeley, CA, 94720-1042.

Malderez, A., & Bodóczky, C. (1999). *Mentor courses: A resource book for trainer trainers*. Cambridge: Cambridge University Press.

McDonough, J., & McDonough, S. (1997). *Research methods for language teachers*. London: Hodder Arnold.

McGrath, I. (2000). Teacher autonomy. In B. Sinclair et al. (Eds.), *Learner autonomy, teacher autonomy: Future directions*. Harlow: Pearson Education.

Mitchell, S. N., Reilly, R. C., & Logue, M. E. (2009). Benefits of collaborative action research for the beginning teacher. *Teaching and Teacher Education, 25*(2), 344–349.

Myhill, D. (2002). Bad boys and good girls? Patterns of interaction and response in whole class teaching. *British Educational Research Journal, 28*(3), 339–352.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge/New York: Cambridge University Press.

Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press.

Orland-Barak, L., & Rachamim, M. (2009). Simultaneous reflections by video in a second-order action research-mentoring model: Lessons for the mentor and the mentee. *Reflective Practice, 10*(5), 601–613.

Oxford, R. (1990). *Language learning strategies: What every teacher should know*. New York: Newbury House.

Patton, M. Q. (1987). *How to use qualitative methods in evaluation*. Newbury Park: Sage.

Polat, B. (2013). Language experience interviews: What can they tell us about individual differences? *System, 41*(1), 70–83.

Porte, G. (2012). *Replication research in applied linguistics*. Cambridge: Cambridge University Press.

Raelin, J. A. (2002). I don't have time to think! versus the art of reflective practice. *Reflections, 4*(1), 66–79.

Rebolledo, P., Smith, R., & Bullock, D. (Eds.). (2016). *Champion teachers: Stories of exploratory action research*. London: The British Council. isbn:9780863558078.

Reid, J. (1987). The learning style preferences of ESL students. *TESOL Quarterly, 21*(1), 87–111. doi:10.2307/3586356.

Richards, J. C., & Burns, A. (2012). *The Cambridge guide to pedagogy and practice in second language teaching*. Cambridge/New York: Cambridge University Press.

Richards, J. C., & Farrell, T. S. C. (2005). *Professional development for language teachers: Strategies for teacher learning*. Ernst Klett Sprachen. Cambridge/New York: Cambridge University Press.

Richards, J., & Farrell, T. (2011). *Practice teaching: A reflective approach*. Cambridge: Cambridge University Press.

Schön, D. (1983). *The reflective practitioner*. New York: Basic Books.

Schön, D. (1988). *Educating the reflective practitioner*. San Francisco: Jossey-Bass.

Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences* (3rd ed.). New York: Teachers College Press.

Smith, R. C. (2000). Starting with ourselves: Teacher-learner autonomy in language learning. In B. Sinclair, I. McGrath, & T. Lamb (Eds.), *Learner autonomy, teacher autonomy: Future directions* (pp. 89–99). London: Longman.

Smith, R. C. (2003). Teacher education for teacher-learner autonomy. In J. Gollin, G. Ferguson, & H. Trappes-Lomax (Eds.), *Symposium for language teacher educators: Papers from three IALS symposia (CD-ROM)*. Edinburgh: IALS, University of Edinburgh. http://homepages.warwick.ac.uk/~elsdr/ Teacher_autonomy.pdf

Smith, R. C. (2015). Exploratory action research as workplan: Why, what and where from? In K. Dikilitaş, R. Smith, & W. Trotman (Eds.), *Teacher-researchers in action* (pp. 235–262). IATEFL: Faversham.

Smith, R., Connolly, T., & Rebolledo, P. (2014). Teacher research as continuing professional development: A project with Chilean secondary school teachers. In D. Hayes (Ed.), *Innovations in the continuing professional development of English language teachers* (pp. 111–129). London: British Council.

Soruç, A., & Griffiths, C. (2015). Identity and the spoken grammar dilemma. *System, 50*, 32–42.

Stenhouse, L. (1975). *An introduction to curriculum research and development*. London: Heinemann.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks: Sage.

Tammenga-Helmantel, M., Arends, E., & Canrinus, E. (2014). The effectiveness of deductive, inductive, implicit and incidental grammatical instruction in second language classrooms. *System, 45*, 198–210.

Tseng, W., Dornyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics, 27*(1), 78–102.

Ullman, R., & Geva, E. (1981). *The target language observation scheme (TALOS): Handbook*. Toronto: Modern Language Center/Ontario Institute for Studies in Education.

Velde, N. (2015). Team teaching for teacher training. In K. Dikilitaş, R. Smith, & W. Trotman (Eds.), *Teacher-researchers in action* (pp. 189–206). IATEFL: Faversham.

Vieira, F., Barbosa, I., Paiva, M., & Fernandes, I. S. (2008). Teacher education towards teacher (and learner) autonomy. In *Learner and teacher autonomy: Concepts, realities, and responses* (Vol. 1, p. 217). Amsterdam: John Benjamins Pub. Co.

Wallace, M. J. (1991). *Training foreign language teachers: A reflective approach*. Cambridge/New York: Cambridge University Press.

Wang, Q., & Zhang, H. (2014). Promoting teacher autonomy through university–school collaborative action research. *Language Teaching Research, 18*(2), 222–241.

Wilson, D. Practical meta-analysis effect size calculator. https://www.campbell-collaboration.org/escalc/html/EffectSizeCalculator-R5.php

Wyatt, M., & Arnold, E. (2012). Video-stimulated recall for mentoring in Omani schools. *International Journal of Mentoring and Coaching in Education, 1*(3), 218–234.

Wyatt, M., & Dikilitaş, K. (2015). English language teachers becoming more efficacious through research engagement at their Turkish university. *Educational Action Research*, 1–21. DOI: 10.1080/09650792.2015.1076731.

Wyatt, M., & Márquez, C. P. (2015). Helping first-year undergraduates engage in language research. *Language Teaching Research*. doi:10.1177/1362168814562013.

Yanati, K. http://yatani.jp/teaching/doku.php?id=hcistats:mannwhitney#effect_size

Yu, X. (2009). A formal criterion for identifying lexical phrases: Implications from a classroom experiment. *System, 37*(4), 689–699.

Yuan, R., & Lee, I. (2015). Action research facilitated by university–school collaboration. *ELT Journal, 69*(1), 1–10.

Index