

Run Randomised Controlled Trials (RCTs)

An introduction

www.nfer.ac.uk



Evidence for Excellence in Education

This document is designed to be read with Adobe Acrobat

Randomised Controlled Trials (RCTs)

This short *How to guide* provides an introduction to Randomised Controlled Trials (RCTs). It is aimed at senior leaders, teachers and other school staff who are interested in carrying out research, with a view to finding out if RCTs are suitable for your school. There is a glossary at the end of this guide which you may find useful to refer to throughout.

What are RCTs?

A Randomised Controlled Trial or RCT seeks to measure impact by identifying a causal link between an intervention and change.

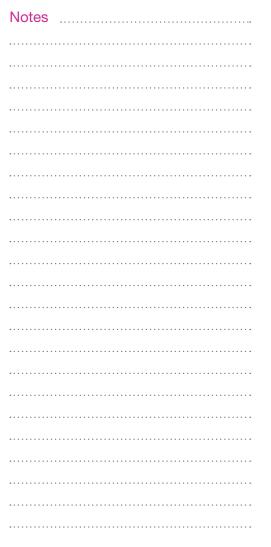
An RCT is a trial carried out on two (or more) groups where participants are randomly assigned to either an **'intervention'** (i.e. the intervention group) or **'controlled condition'** (i.e. the control group). Each group is tested at the end of a trial and the results from the groups are compared to see if the intervention has made a difference – in other words, has the intervention achieved its desired outcome? If the randomised groups are large enough, you can be confident that differences observed are due to the intervention and not some other factor.

RCTs are highly regarded by many organisations involved in education research including the Department for Education (DfE) and the Education Endowment Foundation (EEF). As with any research methodology, RCTs have their place and are not suitable for all research and evaluation.

An RCT is usually run alongside a **process evaluation** of a programme or intervention to see how the intervention has been implemented. This often takes the form of **qualitative interviews** and seeks to find out if the intervention has been carried out as intended. A process evaluation checks the **fidelity of the intervention** (see 'Useful terminology section') and is crucial before assumptions are made about what the results mean. For example, if one teacher is following the intervention in every lesson but another teacher only does the intervention in one lesson a month, this will mean that the groups are not comparable.

To illustrate how an RCT works at the different stages please see Figure 1 on page 3.





'How to' Guides



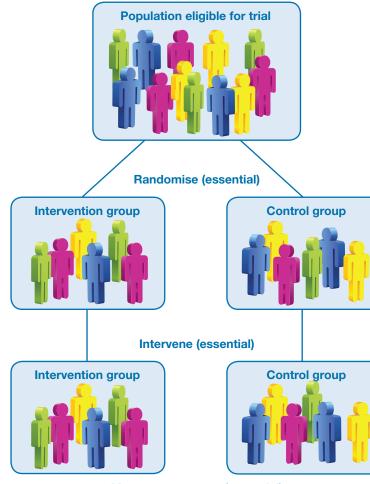
2

Randomised Controlled Trials (RCTs)

3



Figure 1 What is an RCT?



Measure outcomes (essential)

Why do an RCT?

RCTs appeal to schools for a number of reasons:

- They offer a robust methodology that allows you to make causal conclusions. This is because randomisation should mean that, providing they are large enough, the intervention and control groups can be regarded as equivalent and both have a good mix of participants. Differences in outcomes are therefore more likely to be due to the intervention rather than other differences between the groups.
- They are considered one of the best ways of showing if an intervention is achieving its aims and intended impact.
- They can inform school budget decisions.
- There is a need to assess how an intervention is delivered to provide clear guidance for roll-out.



Notes

Ethically and practically, you may want to provide an incentive to your

- We advocate offering control group participants the opportunity to take part in the intervention in the future to ensure all your RCT participants are given the same intervention only at different times. This is not always possible.

Incentives

control group to help encourage them to take part in the research.

There are a number of stages to running an RCT, and several considerations that need to be addressed at the outset.

- Incentives could range from providing additional feedback on performance, to inviting staff to help design the intervention so they are invested in it.

Randomisation

- Randomisation is the deliberate act of randomly assigning participants to either the intervention or control group to ensure their selection to one or other group occurs by chance.
- Randomisation must take place after consent is given.
- It is best to randomise at a pupil level (i.e. randomly assign a learner to one or other group).
- It can be difficult within one school to get enough participants in each group to make sure the trial is big enough to detect the effect (i.e. size of change) we want to see. This is where a pupil-randomised trial can be used across more than one school.

Contamination

- When doing an RCT, contamination must be considered and avoided throughout. An intervention being assessed in a trial should be delivered and work as it would in the real world.
- Bear in mind where contamination could happen, e.g. through learners talking in the playground about what they are doing in class or by staff discussing their work in the staffroom. Even what seems to be an insignificant conversation could result in some of the control group adopting elements of the intervention into their own teaching or learning.
- If there is a risk of contamination (i.e. where the control group can be affected by the intervention group) cluster level randomisation¹ should be used. This is where a whole class takes part in the intervention and a different class is the control group.

Cluster randomisation means that you may need to work with other schools. This also helps ensure you get a big enough sample of classes. While the same stages to undertaking the RCT are required, calculating sample sizes and analysis are more complicated. We therefore recommend you consult an experienced statistician or research organisation to help you.

Randomised Controlled Trials (RCTs)

How to plan an RCT

Figure 2 RCT planning process

Consent

Consent is an essential element

Depending on the intervention being

tested, and/or your school policy, the

headteacher may be able to give

trial is being carried out.

important for fidelity to the

intervention.

consent on behalf of learners. It is

Where school staff are involved in

the trial, they will need to be asked

for consent or at least informed that

the trial is taking place. This is also

good practice to also inform parents

and learners (or participants) that the

Figure 2 outlines them.

for RCTs.

.

•

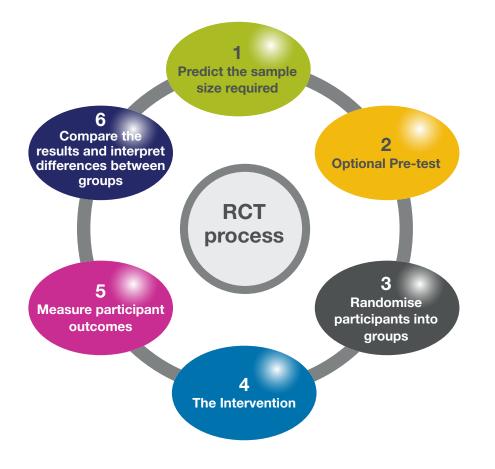
.

4

The RCT process

When carrying out an RCT, there are a number of processes to follow. These are outlined below in Figure 3.

Figure 3 Stages in the RCT process



1. Predict the sample size required

When undertaking an RCT, you must think about the effect you are likely to detect from an intervention – for example, do you expect to see learners scoring an extra ten points on a test as a result of the intervention? If possible, you should look at previous research in the area to inform your thinking and expectations.

Once you have chosen a desirable effect size, you use this to determine the number of participants you need in your trial. Having the right number of participants means you should be able to find an effect (or change) as a result of the intervention, if there is one.

At this stage, you may want or need to work with other schools to give you a larger sample. Having a larger sample means you can detect a smaller effect. However, the larger your sample, the larger the cost implications for the intervention and for your RCT. You will need to consider what is realistic; as with any research, it may be an issue of balancing budget versus effect size.

Educational research is often designed to look at an effect size (see useful terminology section on page 10) of 0.2, which is considered a small effect. However, if you think there will be a larger effect with an effect size of 0.5, then this requires a smaller sample. See Table 1 below for the sample sizes needed to give a small effect and a large effect. You can also alter your effect size between these figures.

Table 1 Effect sizes

Desired effect size	Sample (N)	Intervention group size	Control group size
0.2 (preferable)	800 participants	400 participants	400 participants
0.5	128 participants	64 participants	64 participants

6



2. Optional pre-test (baseline score)

If possible, you should measure your outcome variable for a baseline score (for example, measuring reading ability for a literacy intervention). This needs to happen *before* the randomisation *or* the intervention takes place. Possible ways of doing this include: running a separate pre-test (it is best to use well established commercial tests) or you can use existing data for those participants such as Key Stage 1, 2 or 3 test scores or you could use your own in-school tests.

The test you use at the baseline does not have to be the same as the one you use at the end of the intervention as long as they are highly correlated (i.e. related) to each other. So, for example, you could use a teacher assessment of Kev Stage 3 maths as a baseline measure and then the maths GCSE score after the intervention. Having a baseline score will enable you to measure change and explain some of the variation between scores when you collect outcomes after the intervention. Furthermore, it will allow you to account for any differences or imbalances at baseline in your statistical analysis after the intervention. For example, the control group may have been randomly allocated more learners who eat fruit and vegetables than the learners in the intervention group. Consumption of fruit and vegetables could influence the final outcome measure after the intervention; therefore this would need to be adjusted for in the final analysis of your data.² Using pre-testing

means that the analysis is more sensitive and could also result in reducing the sample size required for the trial. This could make the trial more cost effective and easier to coordinate.

3. Randomise participants into groups

When planning an RCT, you need a plan of action to reduce or eliminate bias. Ideally, randomisation should be carried out by an independent third party, even for small research projects. To learn more about how to randomise please see Hutchison and Styles (2010) available from www.nfer.ac.uk/publications/RCT01. For a 'simple design', when randomising, one group will receive the intervention and the other will carry on as they would have done anyway.

4. The intervention takes place!

This analysis would usually require expert support.

During the intervention it is worth keeping a log of any challenges or issues so that there is an accurate record including where any deviations from the intervention are made (to capture intervention fidelity).

Notes	
•••••	
•••••	 •
•••••	
•••••	
•••••	
•••••	
•••••	



5. Measure participant outcomes

As mentioned previously, this could include using a test that will take place anyway (for example if you are assessing Year 11 learners then you could measure GCSE scores at the end of the year if this is after the intervention has finished) or you could use another test/survey.

It is important that the outcome you are measuring does not relate directly to the intervention. For example, if you are measuring reading then do not use an example about the same topic in the reading intervention **and** the reading test as this gives an unfair advantage. It is best to use pre-developed commercial tests if possible (see www.nfer.ac.uk/nfertests).

6. Compare the results and interpret differences between groups

There are a few factors to consider before comparing and interpreting differences between the intervention and control groups. Firstly, as with any research, make sure you double check your data to ensure there are no errors. Secondly, two types of analysis are normally run in an RCT – 'intention-to-treat' and 'on-treatment'.

Intention-to-treat analysis uses the groups that you randomly allocated participants to before the start of the trial (even though some pupil/classes or schools may have dropped out or not completed the trial).



Notes



On-treatment analysis takes into consideration how the participants have taken part in the trial, for example have they only completed two weeks of a ten week programme?

You will also need to record any missing data. As with any intervention, there is likely to be some drop out which is even more likely within a school setting with learners moving schools and other absences. Therefore, it is best to try and keep everyone engaged with the research from the outset. You must decide what to do with any missing data as there could be different reasons why a participant has not taken part. For example, is it by chance because someone is off sick, or is it deliberate and did everyone who hated maths drop out? These factors may introduce bias and must be accounted for in the analysis. Missing data can require sophisticated analyses and we suggest you contact an experienced research organisation or statistician to help you with this (see Hutchison and Styles, 2010 for more detail on missing data). Once missing data has been accounted for, you can then examine whether the intervention group is statistically significantly different from the control group.

If you **do** have a pre-test, you can add this into your analysis at this point to account for any pre-existing differences in the sample. You will need a statistician or researcher to help you here. If you **do not** have pretest data, you can still compare the difference between the intervention and control groups on the outcome measures. As the samples are random, any differences between the groups should be because of the trial, providing the groups are large enough. At this point you can also examine your intervention fidelity by looking at the log you kept of any issues that occurred during the intervention. This can then be incorporated into the 'ontreatment' analysis.

Figure 4 on the following page outlines some of the advantages and challenges associated with RCTs.



Notes

9

Figure 4 Advantages and challenges associated with RCTs

Advantages

RCTs reduce bias by rendering the intervention and control group equivalent. Bias can be introduced in research in many ways. For example, there is selection bias; (i.e. who is chosen to have the intervention or not). RCTs eliminate this as participants are randomly assigned to one or other group with an equal chance of selection for each.

Random selection is the only way to make causal conclusions. For example, research gives us the ability to say, "Eating five portions of fruit and vegetables a day will increase learners' concentration in class", and not that "Learners who have higher levels of concentration tend to eat more fruit and vegetables". An RCT cannot be used to a) evaluate a policy change that affects many parts of the education system or b) measure the impact of a programme that is rolled out at different times in different places. In these examples, there are too many other factors that could influence change and these factors cannot be controlled for (i.e. you cannot guarantee they will remain the same between the two groups).

Challenges

Some people believe RCTs to be unethical because one group is getting an intervention while another is not. This can be particularly challenging when conducting an RCT on an intervention aimed at underachieving learners. It is important to bear in mind that before you run an RCT, you do not know if an intervention really has a positive impact or whether it has a detrimental effect (this is called the 'equipoise principle'). The purpose of an RCT is to test for these things.

Once you have run a trial you know if it works with that group of participants only (this is called 'limited generalisability') but you must think about who else this intervention would or could work with. It is important to remember that it is better to get the right answer for a smaller group than the wrong answer for lots of people so running an evaluation on a larger sample without randomisation does not solve the problem.

It can be difficult to run blind randomisations in educational RCTs because it is obvious to participants that the intervention is taking place. For example, you cannot create a placebo intervention as you would in medical trials. The intervention group could, therefore, be influenced by the 'Hawthorne Effect'. This occurs when someone knows they are taking part in research so they try harder. To partially overcome this challenge, it is possible to blind test (see 'Useful terminology section') administrators (for the pre- and post-tests), statisticians and in some cases the learners you work with which makes the research more reliable. For example, test administrators are not told which group is receiving which test or statisticians decide on their methods before they see the data so bias is not introduced in any way.

Notes

Useful terminology

Active consent – a type of consent that requires participants (or their parents/carers) to opt-in to a research trial.

Baseline – the point at which all participants start before they have received the intervention. Sometimes a test is given at this point to assess initial differences between the intervention and control groups.

Blinding – this is where someone involved in a trial does not know which group participants belong to (control or intervention).

Blind randomisation – the allocation of participants to either the control or intervention group by someone that is not aware which group is which. Without blind randomisation you can get randomisation bias which happens when participants are allocated to groups in a biased way. For example, because a teacher feels that a learner may need extra help with their reading, they allocate them to the intervention group in a literacy trial.

Cluster level randomisation – this type of trial randomly allocates groups or 'clusters' of participants into either the intervention or control groups. This could be classes within schools or whole schools.

Consent – a participant or someone acting on that participant's behalf agreeing for themselves (or others) to take part in a trial.

Contamination – where the control group starts to run part of (or the whole) intervention.

Control group – the participants in a trial who have been randomly allocated to a group that continues as normal and does not receive the intervention.

Effect size – the size of the difference you are looking to measure with the intervention (statistically speaking this is the fraction of the overall variation that a particular difference represents).

Equipoise principle – the idea that when running a trial we do not know whether it will work and therefore it does not discriminate against any participants depending on what group they are allocated to.

Fidelity – whether an intervention adheres to the guidelines set for the trial.

Generalisability – being able to say that what has been found would also apply to a wider population.

Hawthorne effect – when participants in a trial try harder because they know they are being assessed.

Intention-to-treat – a type of analysis that is entirely based on the groups that participants were randomised to and does not consider whether they actually completed the intervention or not.

Intervention group – the participants in a trial who have been randomly allocated to a group that receives the intervention.

Missing data – when a participant has not completed a test/ survey either partially or fully.

On-treatment – a type of analysis that examines what the groups of participants actually did, for example how many lessons an intervention participant completed of the intervention.

Outcome – this is what the intervention is aiming to produce and could be something like an increase by ten points on a reading test.

Passive consent – a type of consent that requires participants (or their parents/carers) to opt-out of a research trial.

Population – this is the whole group that the sample for the trial can be taken from.

Process evaluation – the side of research that looks at how the intervention has worked or not. It also examines how closely participants have kept to the intended intervention. A process evaluation normally uses qualitative interviews with key stakeholders.

Qualitative data – qualitative research is 'interested in understanding the meaning people have constructed, that is, how people make sense of their world and the experiences they have in the world.' (Merriam, 2009).

Selection bias – this is where participants are chosen to take part in the trial because of a biased reason, for example, they are in a particularly disruptive class at school.

11

Resource considerations

When running an RCT within school there are a few things you need to consider:

- An RCT tells you if something works but not how. Therefore you still need qualitative data, such as interviews or focus groups, to find out what the mechanisms behind the intervention are (i.e. how it works) and how the intervention has led to changes (i.e. what it is about this intervention that makes the difference) through a process evaluation.
- You are **likely to need more than one school** to ensure your research is as robust as possible.
- You will need to let parents know that their child is being involved in an RCT and ask for passive (optout) or active (opt-in) **consent**. Bear in mind, asking for opt-in consent may take more time but if you ask for passive consent, you need to be sure the parents receive the letter.
- You may want to think about holding a parents evening to discuss an RCT and its principles so that misconceptions and concerns about ethics can be explained to parents and staff properly.

Sharing your research

As with any research, it is important to share your results with participants and stakeholders. This will let them know how their involvement has contributed to the research and will also let them know whether the intervention has a positive effect or not. You may need to be careful with how you present these messages if the intervention has shown a negative impact. Offering parents the opportunity to attend a workshop or presentation, for example, may help to alleviate any fears or concerns and will enable them to ask questions.

As RCTs are becoming increasingly popular within the education sector, you may want to share your findings more widely. For example, this may be with other schools, local authorities and/or national research or educational organisations.

For further information about writing up your research see www.nfer.ac.uk/publications/RESM05

Research ideas

Before you decide to test an idea using a trial, please consult the existing evidence to see if it has already been evaluated. One way of doing this is through the Education Endowment Foundation Teaching and Learning Toolkit (www.educationendowmentfoundation. org.uk/toolkit).

Other useful resources

We hope that this short guide to running randomised controlled trials has whetted your appetite for carrying out your own research. NFER has published a series of 'How to' guides for practitioners who want to carry out their own research, helping you put your ideas into practice. NFER have research books and training days available as well as free guidance on topics to research and methods of research. Why not get recognition for your achievements in research in your school, college or early years setting by applying for the NFER **Research Mark? Visit** www.nfer.ac.uk/ris for more information.

References

Hutchison, D. and Styles, B. (2010). A Guide to Running Randomised Controlled Trials for Educational Researchers. Slough: NFER. Merriam, S. (2009). Qualitative Research: a Guide to Design and Implementation. San Francisco, CA: Jossey-Bass.

Randomised Controlled Trials (RCTs)

'How to' Guides



The NFER 'How to' guides are a quick and easy way to digest different aspects of research.

Written by NFER researchers, these guides will help practitioners run research projects in education. From definitions and benefits, through to potential pitfalls, they will ensure the research is based on professional guidance.



© 2014 National Foundation for Educational Research

ISBN 978-1-910008-16-4

How to cite this publication: National Foundation for Educational Research (2014). *How to... Run Randomised Controlled Trials (RCTs): An introduction* (How to Guides). Slough: NFER.

All rights reserved. No part of this document may be reproduced or transmitted in any form or by any means, electronic, mechanical, photocopying, or otherwise, without prior written permission of NFER.

National Foundation for Educational Research

The Mere, Upton Park, Slough, Berks, SL1 2DQ

T +44 (0)1753 637007 F +44 (0)1753 790114 E products@nfer.ac.uk

www.nfer.ac.uk

This guide was published in 2014 and was correct at the time of publication. Users are encouraged to check for the latest advice on data protection with the provisions of the General Data Protection Regulation. For further information please visit the ICO website.

