

Summary

Contents

< [BACK TO TABLE OF CONTENTS](#)

Experiments

By: [Michael Atkinson](#)

In: [Key Concepts in Sport and Exercise Research Methods](#)

Chapter DOI: <https://doi.org/10.4135/9781446288092.n14>

Subject: [Sport & Leisure Studies](#)

Keywords: [baths](#); [control groups](#); [experimental designs](#); [experiments](#); [ice](#); [pilot studies](#); [randomisation](#)

 [Show page numbers](#)

Experiments

My 'running friends' tell me ice bathing is a wonderful recovery method for managing the effects of long bouts of intensive exercise. I hate ice bathing. I cannot imagine a pursuit less pleasurable than lying in a pool of frigid water for 10–20 minutes. But does it really work? If it did, why wouldn't every serious athlete ice bathe on a regular basis? In fact, most of the people I know finish their workout routines by immediately jumping into a hot shower rather than an ice bath! Tips, tricks, techniques and suggestions for training and recovery abound in sport and exercise subcultures. The scientific basis for many of these tips and regimens is practically non-existent or founded on the weakest of empirical grounds. So how do we sort out the rubbish from the sage advice? One of the ways is through the scientific method of experimentation. Consider this example. To test the theory that ice bathing reduces recovery time after long-duration (120 minutes and over) training at a minimum of 75% VO₂ max, I get an idea for research. If the ice bathing theory is correct, then we should find significant improvement in 5km time trial times 48 hours after ice bathing. We could test that out! We have people run in my lab on treadmills on Friday for two hours at 75%, rest Saturday, and then come in for a 5km time trial on a treadmill on Sunday. We do this twice. In the first trial, they have no ice bath, and in the second they have an ice bath. When research is conducted in this manner, we have adopted the thinking and practice of an experimenter.

What Is This Concept?

Experiments are a collection of research designs preferred by sport and exercise physiologists, nutritionists, biomechanists, behaviouralists and many psychologists that use the manipulation and controlled testing of variables to potentially better understand causal processes between them as predicted by theory and stated by research hypotheses. In an experimental research design, one or more variables are manipulated (also called 'treatment' variables) and examined in controlled conditions to determine their effect on a dependent variable (also called 'outcome' variables). Experiments are part of the measurement phase in the scientific method that arbitrates between competing models or hypotheses about a given phenomenon. But where do our ideas for experiments come from? That's very easy to answer ... from theoretical ideas, past research, or previous findings. As described by innovators and proponents of experimental protocols including Galileo, Francis Bacon and Louis Pasteur, experimentation is generally used to test existing theories that purport to explain the causally determined nature of events, entities or processes in the world through the process of deduction. Students need to learn and remember that while we often become caught up in the substantive topics under investigation, the real aim of conducting experiments is to test hypotheses derived from theories.



The most valued kinds of experiments scientists run are called *controlled experiments* (versus 'natural' experiments that occur haphazardly in the world outside of the laboratory). A controlled experiment generally compares the results of a test we design and carry out among an experimental sample against a control sample, which is practically identical to the experimental sample except for the one aspect whose effect is being tested (the independent variable). A good example would be an ice bath experimental trial to see if the method actually works at improving time to full recovery! There are many biological theories justifying why intense cold helps the body to recover after long bouts of exercise. We could organise a simple experiment to test the effectiveness of ice baths. In our experiment, the sample or group receiving ice baths after a 20-mile run would be called the experimental group – they receive the treatment. In order to test whether or not the ice baths are better than no treatment at all, we would need to compare their time to recovery (let's not worry about how to measure this right now) with a group of people who ran the same distance but who did not receive an ice bath afterwards (a smelly group, for sure). The second group would be called our control group. This simple but powerful research design is the basis for practically all experiment protocols and the benchmark design of the scientific method.

The nature of 'control' in experimental designs refers to another matter entirely. Imagine that we actually seek to test the ice bath theory. The very first consideration I need to have in designing the experiment is whether or not any variables I am not measuring or interested in could potentially influence my data if left unaccounted for in the experiment. What if I accidentally included all the young runners in the first group and the veterans in the second? What if one group ran 20 miles of hills and the other 20 miles of a flat course? What if one group ate and drank as recommended for runners during training, and the other starved themselves? Any systematic differences (variables) unaccounted for and unmeasured by experimenters wreak havoc in our studies. We call these pesky variables *confounds* or *extraneous* variables. So, a second aspect of control refers to making sure each of the two groups (in this case) under investigation 'look like' one another in every possible variable we can control other than the treatment variable. This occurs through the process of *randomisation of subjects*.

Randomisation is a process that assigns research participants by 'chance', rather than by choice, to either the treatment group or the control group of all experiments. That is, each study participant has a fair and equal chance of being placed in the treatment group, or being placed in the control group. The goal of randomisation in our example is to produce comparable groups in terms of general participant characteristics, such as age, gender, career history with running, injury background, training volume per week ... and other key factors that affect the probable course of recovery for the runners. In this way, the two groups are as similar as possible at the start of the study. At the end of the study, if one group has a better outcome than the other, the investigators will be able to conclude with some confidence that one intervention is better than the other. A randomised, controlled experiment is considered the most reliable and impartial method of determining what treatments or interventions work the best. Randomisation is normally undertaken in one of two manners:

- *Single-blinded* experiments are those in which participants do not know which group they are in – and therefore which intervention they are receiving until the conclusion of the study.
- *Double-blinded* experiments are those in which neither the participants nor the investigators know to which group the participants have been assigned until the conclusion of the study.

In sum, then, controlled experiments are valued because they must be performed when it is difficult to exactly control all the conditions or variables potentially affecting the outcome of an experiment. Double-blinding, in particular, helps to ensure that the experiment begins by creating two or more sample groups that are *probabilistically equivalent*, which means that measurements of traits should be similar among the groups and that the groups should respond in the same manner if given the same treatment. This equivalency is determined by statistical methods that take into account the amount of variation between individuals and the number of individuals in each group (too complex to cover in this entry). Once equivalent groups have been formed, the experimenter tries to treat them identically except for the one treatment *variable* that he or she wishes to isolate.

Why Is This Relevant to Me?

Asking a scientist why experiments are important in the research process is like asking an ice hockey player why it is important to skate, or swimmers about the importance of water in their sport. But, if an experimenter is pressed to answer the question of relevance, one might expect several of the following answers (aside from their ubiquity across all 'science' disciplines in the university):

- Experiments are very powerful research designs, because if controls are undertaken properly, we can potentially observe cause and effect relationships unfold before our eyes.
- Experiments can be (in most cases) easily replicated by others.
- If experimental findings are replicated across time and space, it adds to the external validity of the findings.



- If the above is true, the method is perfectly suited for testing theory and developing stronger empirical claims regarding causality.

Amongst the 'unwritten' cultural reasons as to why experiments are so fundamental in academic inquiry is that in the modern research-oriented university, the method is practically synonymous with science, research and knowledge. Experiment-based studies command much attention, respect, and indeed funding, if they take as their aim a major problem, theory or dominant belief.

Show Me How It's Used!

What can be overwhelming about experimental designs is that there are so many from which to choose when conducting research. Experiments vary in accordance, for example, with how many hypotheses are being tested, the number of variables involved, how many times the variables are measured before, during and after the treatment is administered, the number of groups and sub-groups in the study, and the duration of the experiment. Phew, that sounds complicated. For the sake of ease, the common types of laboratory experiments employed in sport, health and exercise sciences are outlined below.

Pre-experimental designs – the distinguishing feature of this first category of designs is that it contains no randomly assigned control group, and so extraneous variables pose serious threats to validity. Examples of pre-experimental designs include:

- *One-shot case study*: A treatment is given to one group (an ice bath after 20 miles of running) and we measure time to recovery ... that's it!
- *One-group pre-test, post-test design*: Identical to the above, except that we would have measured time to recovery for the group following a no ice bath scenario on one occasion after a 20-mile run a few weeks before the run/ice bath test in order to get 'baseline' data on the group.
- *Static comparison design*: A design wherein there is a treatment and control group (but not randomly assigned) and each are measured 'post-test' only.

True experimental designs – in these designs, treatment and control groups are employed, and there is random assignment across both groups. Examples include:

- *Post-test only*: Like the static comparison design described above, but with random assignment.
- *Pre-test, post-test, control group design*: The flagship of the experimental protocol. Here, there is a control and treatment group, both are randomly assigned, and measurements are taken on both groups before and after the treatment is administered to the experimental group.
- *Solomon four group design* – An amalgamation of the post-test only and pre-test, post-test designs leading to four groups being involved in the study.

Quasi-experimental designs – these designs are characterised by a lack of random assignment to both the treatment and control groups. Examples of quasi-experimental designs include:

- *Interrupted time series design* – multiple measurements are taken on a dependent variable (time to recovery, for example) over a set period of time, both before and after a treatment is introduced to the experimental group. In this scenario, for example, we might study runners over nine months, asking them each to run 20 miles on the last Sunday of each month. For the first four months the participants do nothing after their runs, but at month five we administer the ice bath to the treatment group. Then for the next four months, no further treatment is administered.
- *Non-equivalent control group design* – this looks and 'feels' like the gold standard pre-test, and post-test design, but there is no random assignment of subjects to either group. This occurs quite frequently when researchers have not been able to gather their entire sample group before the study commences.

Whatever the experimental design one selects (and the above are not all of the experimental designs available, only the most common), several steps in the experiment process are worth undertaking before one actually conducts the study.

First, to avoid being forced into performing a quasi-experimental design, one should attempt to obtain a full sample before the research commences. Such is not always feasible due to time constraints, competition for lab space, worries about participant schedules and needs, and a host of other factors. Second, it may be wise to first conduct a very small-scale pilot or test study or two before you do the real experiment. Pilot studies help ensure that the experiment measures what it should, and that everything is set up right in the lab setting. Minor errors, which could potentially destroy the experiment, are often found during the pilot test process. Are my thermometers measuring water temperature accurate? If I am using several ice baths in the lab, and one of the thermometers is off, this translates into systematic error in my design. With a pilot study, you can gather information about potential errors and problems, and improve the design, before putting a lot of effort into the 'main event'. If the experiments involve humans, a common strategy is to first have a pilot study with someone involved in the research protocol itself (good luck getting a volunteer from the research team to run 20 miles and then jump in an ice bath), and then arrange a pilot with a person who resembles the subject(s). Those two different pilots are likely to give the researcher important information about any problems in the experiment.

Problems, Pitfalls and Controversies

The most consistently voiced criticisms of experimental designs focus on threats to internal validity brought about by problems in the design or practice of the method. While staunch defenders of experiments claim internal validity (am I really measuring what I think I am measuring?) as among the strongest selling points of the method, critics are accurate in mentioning a lengthy list of serious threats to internal validity. Researchers undertaking lengthy trials are often unable to control for the *maturation effect*: the condition in which subjects 'naturally' change in their abilities or characteristics due to factors beyond experimental manipulation. Runners get injured outside of the lab, or they might progress and improve their physical conditioning over the course of a year and thus are not affected in the same way by ice bathing. Both situations might dramatically impact the findings in our study. More potentially damaging than maturation effects are *test effects*, in which participants become accustomed to being tested and adjust their behaviours accordingly (perhaps they run very hard in the first test, and then take it easier in the second test). *Experimental attrition* (subjects withdrawing from the study) can alter the balance of equality between the two groups (as established by randomisation) and thus weakens our ability to be sure extraneous variables have been managed sufficiently in the study.

More damning challenges to the integrity of experimental designs come from attacks on the *external validity* of the findings. While the external (population) significance of sample data can be examined and argued through inferential statistical tests, unless sampling has been carried out through a sufficiently random fashion from the theoretical population, there are scant reasons to believe the data represent the population (thus, there is poor *population validity*). If in my test of ice bathing, only very novice runners volunteer, can I make legitimate claims about ice bathing's impact on runners in general? In a related way, if there is little confidence in the *ecological validity* of the experimental protocol, then researchers are in quite a quandary. Questions of ecological validity focus on whether findings in the lab would hold true in other contexts. If our experimental results can only be obtained in the lab using treadmills and specially constructed baths, but not where people actually live, interact, and do things like running or bathing, then what is the use of the study?

See also: [Causality](#); [Evaluation Research](#); [Evidence-Based Research and Practice](#); [Inferential Statistics](#); [Meta-Analysis](#); [Populations and Samples](#); [Positivism](#); [Variables](#).

Key Readings

Field, A. and Hole, G. (2006) *How to Design and Report Experiments*. Thousand Oaks, CA: Sage.

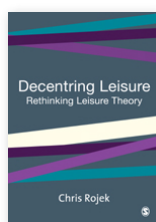
Freedman, D. (2009) *Statistical Models: Theory and Practice*. Cambridge: Cambridge University Press.

Thomas, J. and Nelson, J. (2001) *Research Methods in Physical Activity*. Champaign: Human Kinetics.

< [Evidenced-Based Research and Practice](#)

[Grounded Theory](#) >

Read next in Sage Academic Books



 Academic Books

Books



 Academic Books

Books

Globalization & Football



 Academic Books

Books

Key Concepts in Sports Studies



 Academic Books



Decentring Leisure: Rethinking
Leisure Theory
Rojek, Chris
December 20, 2013

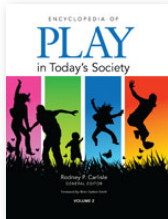
Giulianotti, Richard, et al.
June 19, 2012

Wagg, Stephen, et al.
January 18, 2013

Key Concept
Botterill, David, et
Research
December 22, 201



More like this in Sage Knowledge

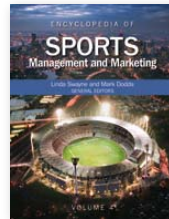


 Reference

Encyclopedias

Encyclopedia of Play in Today's
Society

May 18, 2009



 Reference

Encyclopedias

Encyclopedia of Sports
Management and Marketing

August 25, 2011



 Reference

Handbooks

Handbook of Physical
Education

June 22, 2009





En

The SAGE Di
Studies

Malcolm, Dominic
August 2, 2012



Sage Recommends

We found other relevant content for you on other Sage platforms.



About

[About Sage Knowledge](#)

[About Sage Publishing](#)

[Accessibility](#)

[Author Guidelines AI/LLM](#)

[CCPA](#)

[Permissions](#)

[CCPA – Do Not Sell My Personal Information](#)

[Privacy policy](#)

[Terms of use](#)

Information for

[Authors](#)

[Frequently Asked Questions](#)

[Editors](#)



Librarians
Students and researchers
Trial Login

Stay in Touch

Contact us

Copyright © 2025 by Sage Publications

