

Lab 2

Table of contents

Psychological research basics	3
□ Lab class	3
□ Self-study	5
Empiricism	5
Conceptual and operational definitions	6
Types of claims	6
Reliability and validity	7
Confirmation	9
Research producers vs consumers	10
□ Self-study	10
Confirmation	11
Quiz 1	12
Explore, apply, reflect	14
□ Lab class	14
Lab 2 practice quiz	14
Research basics activity	14
□ Self-study	16
Food for thought: Paracetamol and autism	16
Confirmation	17
References	17

Psychological research basics

“A deeper understanding of judgements and choices also requires a richer vocabulary than is available in everyday language.”

— Daniel Kahneman, *Thinking, Fast and Slow*

□ Studying psychology is in some ways like learning a new language. You need to acquire a vocabulary for expressing psychological facts and ideas. In some cases, this will require learning new terms. In other cases, this will require understanding the specific meaning of more common terms in a psychological context (e.g., significant, reliable, valid).

“Die Grenzen meiner Sprache bedeuten die Grenzen meiner Welt.”

(“The limits of my language mean the limits of my world.”)

— Ludwig Wittgenstein, *Tractatus Logico-Philosophicus*

□ Without mastering this vocabulary, your ability to grasp psychology as a scientific discipline will remain limited.

“Extraordinary claims require extraordinary evidence.”

— Carl Sagan, *Broca's Brain*

□ In psychology, as in other sciences, bold or surprising hypotheses demand strong empirical support. The more a claim challenges established knowledge, the more rigorous the evidence must be before it can be considered credible.

“Habe Mut dich deines eigenen Verstandes zu bedienen.”

(“Have the courage to use your own reason.”)

— Immanuel Kant, *What Is Enlightenment?*

□ Kant argues that true progress, for both individuals and society, only happens when people stop blindly following authority and start thinking for themselves. It's a challenge to question what you're told and form your own conclusions, rather than passively accepting information from others. This intellectual courage is the foundation of all critical thinking and true learning.

Lab class

In this lab, we will cover issues related to Chapters 1 to 3 and 5 in Beth Morling's book. The first three chapters in Beth's book are an introduction to scientific reasoning and Chapter 5 focuses on identifying good measurement. Please note that your knowledge of these chapters will be tested in a summative quiz (see Section).

Some of the key concepts covered in Chapters 1 to 3 and 5 are:

Empiricism

Psychological reasoning must be based on **data**, not opinions or intuitions.

Measured and manipulated variables

A **measured variable** is one whose levels (the different values or categories it can take) occur naturally and are simply observed and recorded by the researcher.

A **manipulated variable** is one whose levels (e.g., experimental conditions such as training vs. no training) are controlled by the researcher. True experiments are based on manipulated variables (the term true is sometimes added to distinguish them from quasi-experiments, which do not involve manipulation).

Some variables can only be measured (e.g., height or intelligence), whereas others could be measured or manipulated. If, say, you are interested in the effect of caffeine consumption on exam performance, you could ask participants how much coffee they drank before the exam (measured) or you could assign them to different levels of caffeine intake before the exam (manipulated).

Conceptual and operational definitions

A **conceptual definition** is a general explanation of what the concept means, based on theory or scholarly definitions. For example, Neisser et al. (1996) proposed the following conceptual definition of intelligence: “[Intelligence is the] ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought.” (Neisser et al., 1996, p. 77)

An **operational definition** on the other hand describes how to measure or observe a concept. They describe what you need to do (i.e., which operations to perform) to measure or observe something.

Types of claims

There are three prototypical types of claims that can be made in the context of psychological studies:

- **Frequency claims:** The aim is to measure a single variable as accurately as possible. Example: “39% of teens admit to texting while driving.”
- **Association claims:** Investigate associations or correlations between variables. Example: “Coffee consumption linked to lower depression in women.”
- **Causal claims:** Claim that one variable (the **independent variable or IV**) causally influences another (the **dependent variable or DV**). Causal claims can only be based on true experiments. Example: “Spatial working memory training improves navigation skills.”

Reliability and validity

Two key concepts for evaluating psychological research are reliability and validity.

Reliability refers to how consistent a certain measurement is.

Validity refers to how well it measures what it is supposed to measure.

A good way to illustrate this is with scales. Imagine you weigh a 1 kg bag of flour on your kitchen scales ten times:

- If you get 10 different weights, the scales are **not reliable**. They are also **not valid**, because a measure that is not reliable cannot be valid.
- If the scales show 800 g every time, the measurement is **reliable** (because it's consistent) but **not valid** (because the true weight is 1 kg).
- Only if you get 1 kg each time is the measurement **both reliable and valid**.

Self-study

If you took A-level psychology, some of the content in Chapters 1-3 and 5 in Beth's book might sound familiar. Nevertheless, we would encourage to carefully read these chapters and compare their content to what you were told in school. Check if there are things you hadn't heard about before. Ask yourself if what you read is consistent with what you previously learnt.

Empiricism

An important point made in Chapter 1 of Beth's book is that psychologists are empiricists. We have pointed out above that psychological reasoning must be based on data, not opinions or intuitions.

Just to be clear: There is nothing wrong with opinions or intuitions. On the contrary. However, they can only ever be the beginning, but not the endpoint of psychological thinking. We need to test our opinions and intuitions empirically, that is, by conducting studies that generate data.

As an empiricist, you should not naively accept claims, but critically question them. Below are some key questions you could ask when encountering a claim.

! Important

- What does this claim actually mean?
- Given what you already know, how credible is the claim?
- What evidence is there to support the claim?

If there appears to be evidence to support the claim:

- Is the explanation provided plausible? What could be alternative explanations for the effect? E.g., does the study have methodological shortcomings?
- Is the effect statistically significant?
- How big is the **effect size**?

Conceptual and operational definitions

When it comes to measuring intelligence, [various operational definitions](#) have been suggested. One example is Raven's Progressive Matrices:

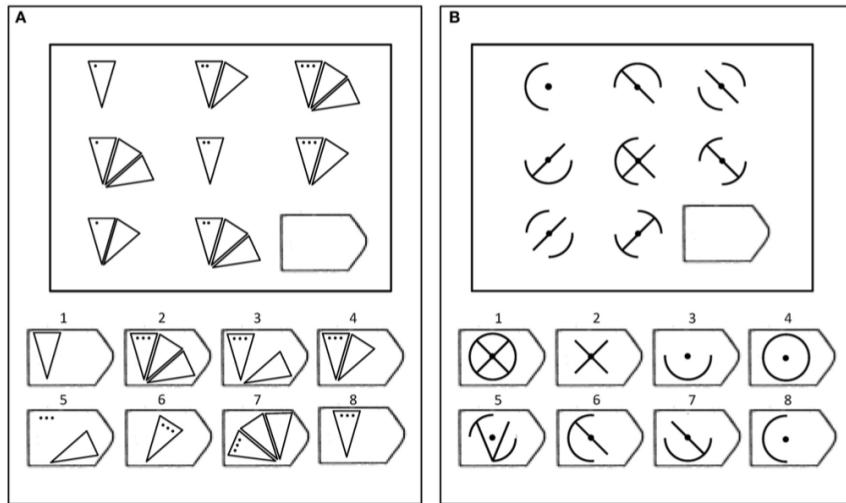


Figure 1: Two examples similar to those used in Raven's test. Participants must identify the item that completes the overall pattern in a rule-based manner. From Little et al. (2014).

Here the operational definition involves presenting participants with complex patterns they have to complete. Working out the correct answer arguably involves reasoning and problem solving. Whether or not this is indeed an appropriate way to measure intelligence is a matter of validity (see below).

Note that **conceptual and operational definitions are not independent**. If, for example, your conceptual definition of intelligence suggests that visuospatial and verbal abilities are separable, your operational definition of intelligence must reflect this distinction. That is, your test must include items that are assumed to measure visuospatial abilities and items that are assumed to measure verbal abilities.

Operationalisation can take rather different forms, each with their own advantages and drawbacks. Imagine you wanted to investigate the frequency of texting while driving (your conceptual variable). You could simply ask people how often they text and drive. These data would be simple to collect, but might suffer from [social desirability bias](#). Alternatively, you could hire research assistants to directly observe drivers in their cars. This would likely give you a better idea of the actual texting frequency in cars, but would be a very time-consuming and costly way of collecting the data. Finally, you could ask people how often *their friends* text and drive. You might hope that this approach removes some of the social desirability bias, while at the same time making the data easy to collect.

Types of claims

Note: The type of data determines the type of claim you can make:

Measured variables □ Association claims

If you asked participants how much coffee they drank before the exam and found that those who drank more coffee performed better, you could make an association claim: “Caffeine intake is associated with better exam performance.”

Why is a causal claim not possible? The association between coffee intake and better exam performance might not be due to the coffee intake per se, but could be explained by a **third, unmeasured variable**: For example, the association between drinking coffee before an exam and better performance could be explained by **personal organisation and time management skills**.

- **The Well-Organised Student:** This student likely has their study schedule managed, their materials prepared, and their morning routine planned. Because they aren't rushing or panicking, they have ample time for a normal morning routine, which may include making and drinking a cup of coffee.
- **The Less-Organised Student:** This student might have crammed the night before, overslept, and is now rushing to get to the exam on time. In their haste, they skip breakfast and have no time to make coffee.

In this scenario, the coffee isn't *causing* the better performance. Instead, the ability to **have time for coffee** is simply a **marker of a well-organised student**. It's the underlying organisational skills which lead to better study habits and a less stressful exam day experience that is the true cause of the better performance.

Manipulated variables □ Causal claims

If you randomly assigned participants to two groups (caffeine vs. placebo), controlled their caffeine intake, and found that those in the caffeine group performed better, you could make a causal claim: “Caffeine intake leads to better exam performance” (at least under the specific conditions you tested)

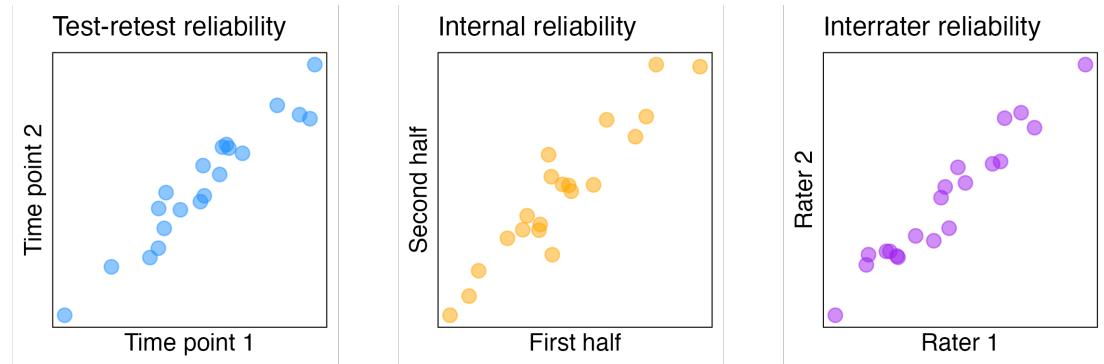
Note that the coffee example is a made-up one to illustrate the difference between claims based on measured and manipulated variables. If you're interested in the actual effects of caffeine on cognition, check out these studies: James (2014), McLellan et al. (2016), Ramirez de la Cruz et al. (2024), Ricupero & Ritter (2024)

Reliability and validity

Reliability can be measured in the following ways:

- **Test-retest reliability:** Can be used for experimental or questionnaire studies. The basic idea is to repeat the study after a delay with the same participants and to investigate how similar the participants performed on both attempts.
- **Internal reliability/consistency:** Typically used for questionnaires. The basic idea here is to split a questionnaire into different parts (e.g., two halves) and to investigate how well these parts correlate with each other.
- **Interrater reliability:** Typically used where independent observers rate certain behaviours. The idea is to investigate how similar the ratings across observers are.

Note that all types of reliability predict a high positive correlation if measurements are reliable:



< fa comments > Activity

Can you explain what the data points in each of the plots represent? Once you have answered this question, ask an AI for the answer. You can do so by describing the plots or by uploading a screenshot of the plots. Do you and the AI agree? If you don't agree and you're not sure who's right, let us know in the next lab!

In Chapter 3, Beth refers to four “big validities”:

1. **Construct validity:** How well has the researcher defined and measured (frequency and association claims) or manipulated (causal claims) the variables of interest? Does the test correlate with other tests that measure related constructs? (□ convergent validity) Does the test not correlate with tests that measure unrelated constructs? (□ discriminant validity)
2. **Statistical validity:** How precise is our estimate and how big is the effect size?
3. **External validity:** How well do the results generalise to different people, times and places?
4. **Internal validity:** To what degree can we be sure that there are no alternative explanations for the results? (relevant for causal claims)

As the scales example above shows, **reliability is necessary but not sufficient for validity**.

It is **necessary** because a highly unreliable test cannot be valid. In fact, a test's reliability sets the upper limit for its validity. That is, a test cannot be more valid than it is reliable.

Why? The highest possible correlation a test can have is with itself (which is a measure of reliability—test-retest reliability). No other test can correlate more strongly with a test than the test itself (and correlations with other tests are a measure of validity—convergent validity).

However, reliability is **not sufficient**, as shown by the scales example: Even a highly reliable measurement is not necessarily valid.

Confirmation

! Important

Please confirm you have worked through this chapter by submitting the corresponding chapter completion form on [Moodle](#).

Research producers vs consumers

Self-study

In my view, Beth Morling makes a very useful distinction between research producers and research consumers. Examples for research producers are researchers at universities, PhD students or undergraduate students conducting research projects. Examples for research consumers are clinical psychologists, educational psychologists, counsellors, organisational psychologists, and many other types of psychologists working in applied fields.

A [Careers Destinations report](#) published by the [British Psychological Society](#) (BPS) showed that only about 7% of psychology graduates work in scientific research and development:

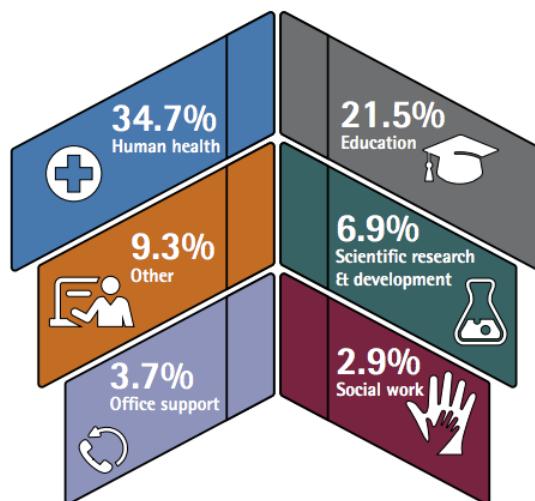


Figure 1: Employment sectors for psychology graduates. Source: Morrison Coulthard, L. J. (2017). *BPS Careers Destinations (Phase 3) Survey 2016 Report*. The British Psychological Society.

If it is somewhat unlikely that you become a research producer, why should you care about research? Beth's and my answer is that it is just as important to understand how research is produced if you are a research consumer. Why? In my view, applied psychology can only be as good as the research that underpins it. As an applied psychologist, you will need to make decisions, for example, you might need to decide what type of therapy to recommend to a client. (Even as a student, friends or family members might ask you for advice!¹)

¹One of my teachers once said that hearing that you study psychology will divide people into two groups: Those who take a step forward, and those who take a step back.

These recommendations should be based on scientific evidence. I would argue that as psychologists, we have a duty to make the best possible recommendations given current scientific evidence. To make these recommendations, a research consumer needs to be able to understand and critically analyse research. And to learn and improve on these skills, it is important to also have produced research. Therefore, we will focus on both critically analysing existing research as well as producing our own research in this module.

Confirmation

! Important

Please confirm you have worked through this chapter by submitting the corresponding chapter completion form on [Moodle](#).

Quiz 1

Quiz details

- Available from **Thursday, 9 October at 3pm** in the **Quizzes and assignments** section on Moodle.
- **Deadline: Thursday, 16 October, 3pm.**
- There will be 53 questions about Chapters 1-3 and 5 of Beth's book.
- All questions will be four-alternative multiple choice, with no penalty for incorrect answers.
- You will receive a mark corresponding to the percentage of correctly answered questions. For example, if you get 40 out of 53 questions correct, your mark will be 75.
- The quiz is time-limited: Once started, you have 50 minutes to complete the quiz in one sitting (therefore, make sure to read the book chapters *before* attempting the quiz; just to be clear: you may use the book to look things up—it just wouldn't be a good strategy to start reading the book when you start the quiz...).
- You have one attempt at the quiz and you will not be able to retake the quiz.
- This is one of the three quizzes that together contribute 10% to your overall module mark (see [?@sec-assessment](#)).

If you have a support plan:

- The time limit can be adjusted in accordance with your support plan. Please [contact me by email](#) well before the deadline if you require an adjustment of the 50-minute time limit.
- If on the other hand you need a deadline extension, please fill in the [coursework extension form for students with support plans](#).

Use of AI

Table 1: AI Use Guidelines for the Assessment

Key academic activities for this assessment	AI use	What you could use AI for	AI use documentation	Where you can learn to do this without AI
Read and understand the key concepts in the book chapters	Optional	Explain concepts. Generate illustrative examples. Generate questions.	Not required	Complete end-of-chapter activities in the book. Discuss concepts with peers. Ask questions in the lab or on the forum.

Complete the quiz

Prohib-
ited

Be aware that outputs of generative AI are not peer-reviewed, might include incorrect made-up information (hallucinations) and may omit key information and viewpoints.

Quiz FAQs

Q: Can I interrupt the quiz?

A: *In theory*, it should be possible to interrupt a quiz and return to it later. The [Moodle help pages](#) state: “If you answer a number of questions and then want to complete the quiz at a later time, click ‘Finish attempt’. When you return to continue the quiz, Moodle will remember which page you were on and allow you to continue from there, having saved your previous questions.”

In practice, students occasionally have had issues with returning to a quiz in the past, which is why we recommend to complete the quiz in one sitting (i.e., without interruption). Of course, you can still open other browser windows or tabs while working on the quiz.

If it turns out an interruption is unavoidable, follow the advice from the Moodle documentation and let us know if you encountered any issues when returning to the quiz.

Q: Do I have to check if my answer was correct?

A: No. You can, but you do not have to check the accuracy of your answers (i.e., you can move on to the next question without checking the accuracy and your answer will still be saved). If you have questions about the quiz, please ask on the Moodle forum before starting the quiz. Upon completion of the quiz, you will be shown your overall quiz mark.

Q: How difficult will this quiz be and how should I prepare for it?

A: The main purpose of the quiz is to ensure that you have read the relevant book chapters and have understood the key concepts. To give you an idea of the quiz difficulty, over the past few years, the average quiz mark has ranged between 75 and 80%, and the percentage of fail marks (i.e., marks below 40%) for this quiz has ranged between 1 and 3%.

If you like to create an excerpt of the material you read, feel free to do so. While probably not strictly required for the quiz, taking notes now will likely help you when it comes to revising for the exam in January.

Explore, apply, reflect

Lab class

Lab 2 practice quiz

Here is a short, unmarked practice quiz to help you check your understanding of the key concepts we've covered. It's just a few questions and is designed for your own learning.

[Link to Practice Quiz](#)

We'll go over the answers once you've completed it.

Research basics activity

This activity builds on the psychological questions we generated last week:

- [Whiteboard Group 1 \(Tue, 9-10:30am\)](#)
- [Whiteboard Group 2 \(Tue, 10:30am-12pm\)](#)
- [Whiteboard Group 3 \(Wed, 9-10:30am\)](#)
- [Whiteboard Group 4 \(Wed, 10:30am-12pm\)](#)
- [Whiteboard Group 5 \(Thu, 9-10:30am\)](#)
- [Whiteboard Group 6 \(Thu, 10:30am-12pm\)](#)

Goal: To practice thinking like a psychologist by taking a general question and outlining a simple research study to investigate it.

Instructions:

1. If possible, work in a pair or a small group of up to three. Open the whiteboard from last week.
2. Choose **one question** to investigate. If your chosen question is very broad (e.g., "What makes us happy?"), your first step is to narrow it down into a more specific, **testable research question** (e.g., "Does spending 30 minutes outdoors each day increase self-reported happiness levels?").
3. Open a Word document and write down your refined research question. Then, answer the following points.
 1. **Study Design:** Propose a simple design. Would you conduct an **experimental study** (where you manipulate a variable) or a **correlational study** (where you just measure existing variables)? Briefly explain why.
 2. **Variables:** What are the key variables in your design? For each one, state whether it is **manipulated** or **measured**.

3. **Operational Definitions:** How would you operationally define each variable? That is, how would you turn the abstract concept into a concrete, measurable thing?
4. **Type of Claim:** Based on your design, what is the strongest type of claim you could make? An **association claim** (that two things are related) or a **causal claim** (that one thing causes another)?

Class Presentation: Time permitting, we will ask some groups to briefly present their answers in class. Be prepared to share your research question and your thinking on a few of the points above.

 Tip

Stuck? Ask for Help! You can ask us for guidance, or you can use [Copilot](#). If you use Copilot, try specific prompts like “Help me operationally define ‘social anxiety’ for a survey study.” If you present your answers in class and you used Copilot, please explain how you used it and how helpful you think the answers were.

Click here to see a worked example

Here is a worked example to illustrate how the activity could be completed:

- **Original Question:** “Does coffee make you smarter?”
- **Refined Research Question:** “Does drinking one cup of coffee improve performance on a simple memory test?”
- **Design:** An **experiment**. We manipulate coffee intake to see if it *causes* a change in memory.
- **Variables:**
 - Caffeine Intake (coffee vs. no coffee): **Manipulated**
 - Memory Performance (score on test): **Measured**
- **Operational Definitions:**
 - “Caffeine Intake” = Participants drink either a cup of standard brewed coffee or a cup of decaf coffee (placebo).
 - “Memory Performance” = The number of words (out of 20) a participant correctly recalls 5 minutes after studying a list.
- **Claim:** Because this is an experiment, we could make a **causal claim** if the coffee group performs significantly better.

If you finish early, here are some optional questions for you to consider:

- **Validity & Reliability:** What is one step you would take to improve the **validity** of your study (i.e., ensure you’re measuring what you intend to)? What is one step you would take to improve the **reliability** of your study?
- **Sampling:** Who would you recruit as participants for your study? Why did you choose this group?
- **Ethics:** What is one potential ethical issue you would need to consider?

Self-study

Food for thought: Paracetamol and autism

As mentioned earlier, the coffee example was hypothetical. Here, however, is a real case with potentially serious consequences. On 22 September 2025, President Trump advised pregnant women not to take paracetamol, claiming it could increase the likelihood of their child developing autism.

To be clear, the causes of autism are complex and not fully understood. In an [article in the New York Times](#), Alison Singer, the president of the Autism Science Foundation, noted:

Autism doesn't have a single cause. It is the result of a complex mix of genetics and environmental factors. We know that genetic factors play the biggest role; hundreds of genes have been linked to autism, and inherited or spontaneous changes in these genes can alter brain development. Environmental factors also matter, especially during pregnancy, such as advanced parental age at conception, prematurity or low birth weight, and exposures that affect brain development, like fever or illness during pregnancy.

In fact, there could be an association with paracetamol as Brian Lee, an epidemiologist, explained in the same article:

There are a number of studies, including our [2024 study in JAMA](#), that showed an apparent statistical association between [paracetamol] use during pregnancy and children's risk of autism, A.D.H.D. and intellectual disability. But association is not causation.

Note the key phrase "**association is not causation**". This raises the question of whether third variables might explain both paracetamol use and increased autism risk. Alison Singer pointed out:

The key question is: Why are these pregnant women taking [paracetamol] in the first place? We know that fever during pregnancy is a risk factor for autism. So if they were taking [paracetamol], was it the fever that caused the autism or the [paracetamol]?

This illustrates how third variables can complicate causal interpretations: infections during pregnancy can cause fever, which leads women to take paracetamol, but those same infections might themselves increase the risk of autism. In other words, **the observed association between paracetamol use and autism could be driven by a third variable** (the infection) rather than the medication itself.

Further thought: If you were designing a study to test whether paracetamol use during pregnancy causes autism, how would you address the possibility of third variables like infections or fever?

Think, then click here to read this model answer

One way to deal with third variables is to collect information about them and take them into account when analysing the data. For example, if fever or infections could be the real reason for the increased autism risk, researchers should record whether the mothers had infections during pregnancy. Then they can compare groups that are similar

in terms of infections but differ in paracetamol use. Another idea is to use a study design that reduces differences between groups, like comparing siblings where one pregnancy involved paracetamol use and the other didn't. This helps because many other factors (like genetics) stay the same. However, designing such studies is challenging because it's hard to measure all possible third variables accurately, and the study would need to follow participants for years - from pregnancy through early childhood - to see whether autism develops. This is why many studies of this type are retrospective, using data collected after the fact. But that creates another problem: mothers have to remember whether they had an infection and whether they took paracetamol during pregnancy - which is also quite challenging!

If you're interested in the topic, I would also recommend to read the following articles:

- [Should the Autism Spectrum Be Split Apart? \(The New York Times\)](#)
- [How the White House Spun 'Weak' and 'Inconclusive' Studies to Tie Tylenol to Autism \(The Guardian\)](#)
- [Let's Talk About What Autism Actually Is \(The New York Times\)](#)
- [Autism Should Not Be Seen as Single Condition With One Cause, Say Scientists \(The Guardian\)](#)

Confirmation

! Important

Please confirm you have worked through this chapter by submitting the corresponding chapter completion form on [Moodle](#).

References

James, J. E. (2014). Caffeine and cognitive performance: Persistent methodological challenges in caffeine research. *Pharmacology Biochemistry and Behavior*, 124, 117–122. <https://doi.org/10.1016/j.pbb.2014.05.019>

McLellan, T. M., Caldwell, J. A., & Lieberman, H. R. (2016). A review of caffeine's effects on cognitive, physical and occupational performance. *Neuroscience and Biobehavioral Reviews*, 71, 294–312. <https://doi.org/10.1016/j.neubiorev.2016.09.001>

Neisser, U., Boodoo, G., Bouchard, T. J. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *The American Psychologist*, 51(2), 77–101. <https://doi.org/10.1037/0003-066X.51.2.77>

Ramirez de la Cruz, M., Esteban-Garcia, P., Abian, P., Bravo-Sanchez, A., Pinas-Bonilla, I., & Abian-Vicen, J. (2024). Effects of different doses of caffeine on cognitive performance in healthy physically active individuals. *European Journal of Nutrition*, 63(8), 3025–3035. <https://doi.org/10.1007/s00394-024-03486-9>

Ricupero, S., & Ritter, F. E. (2024). Caffeine and cognition: A cognitive architecture-based review. *Theoretical Issues in Ergonomics Science*, 25(6), 655–679. <https://doi.org/10.1080/1463922X.2024.2323547>