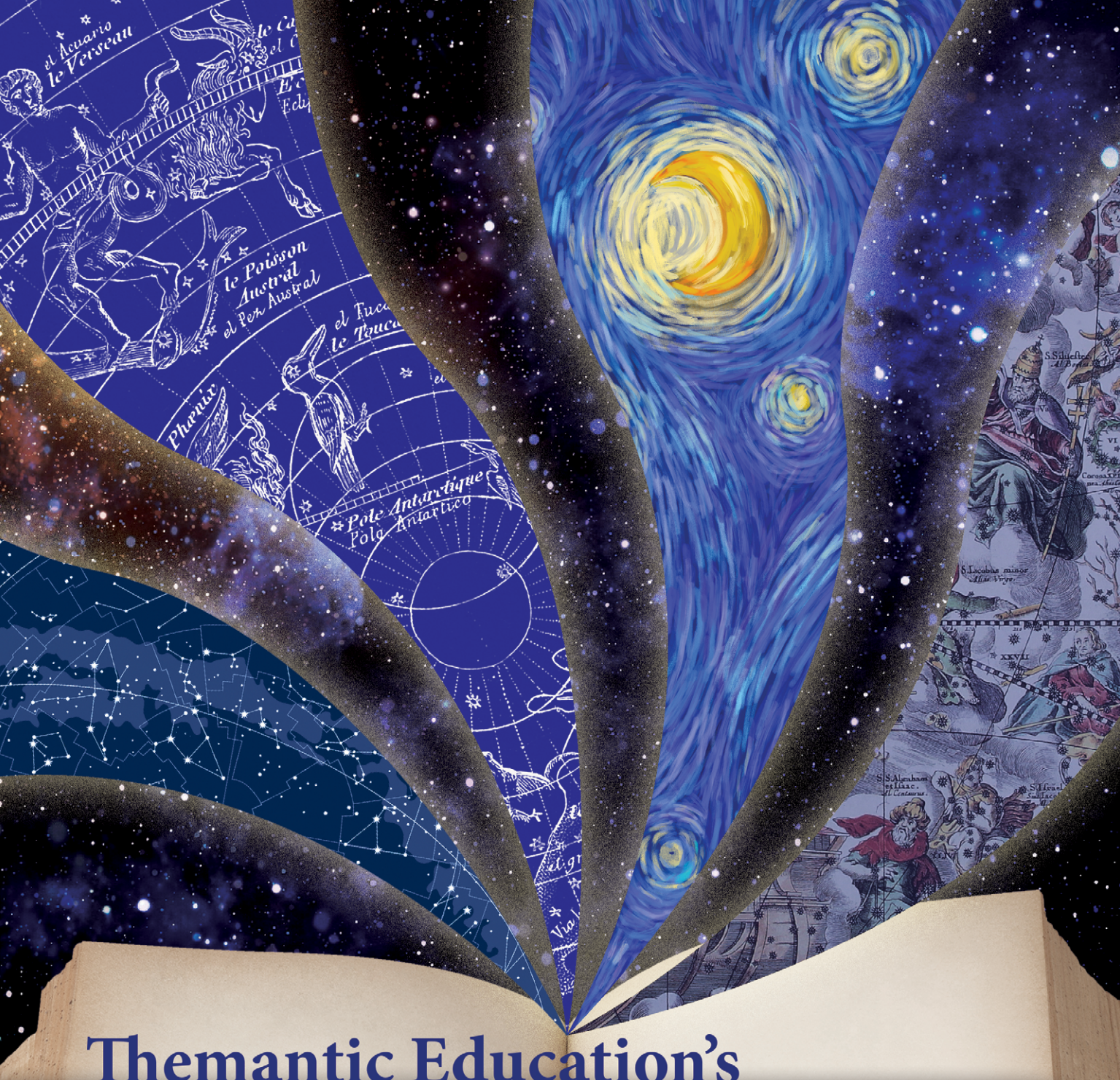


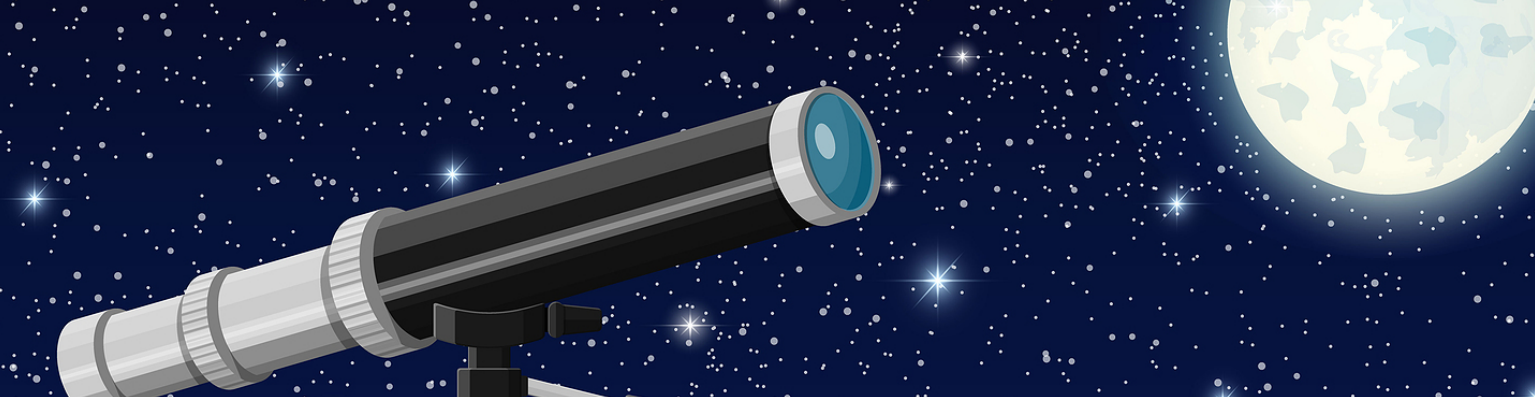
Themantic Education's

IB Theory of Knowledge

A Student's Guide

PREVIEW

Alexey Popov



CONTENTS

	Introduction	7
Unit 1	Knowledge of knowledge	25
Unit 2	Knowledge and technology	71
Unit 3	Bias in personal knowledge	153
Unit 4	Bias in shared knowledge	199
Unit 5	Knowledge and understanding	309
Unit 6	Knowledge and language	399
Unit 7	Assessment guidance	481
	Glossary	519
	References	539

HOW TO USE THIS BOOK

If you are a Theory of Knowledge student, this book is for you. It is designed to be used in class or at home as a student's guide to the IB TOK course. Here is an overview of the features that you will find here.

Lessons

The book is broken down into lessons. Each lesson includes the following elements:

1. **Learning outcomes**. These are key guiding questions that you will be able to answer at the end of the lesson. There belong to three levels:
    a. *Knowledge and comprehension*: this is about knowing the key concepts or ideas and being able to explain their meaning
    b. *Understanding and application*: this is about being able to apply the concepts to specific scenarios or problems, and also to see how different ideas link to each other
    c. *Thinking in the abstract*: this is about understanding some abstract, often debatable problems of knowledge in general
2. **Key concepts**. Usually every lesson is focused on one key knowledge concept (for example, doubt, justification, bias). Sometimes there are a few other concepts that are closely related to this central one. In the lesson itself all key concepts are printed in red font. If you see the red font, it means that the concept is included in the Glossary at the end of the book.
3. **Other concepts used**. These are concepts that are being discussed in the lesson, but are not central to your understanding of knowledge problems. Usually these concepts are related to some specific theories or examples that are used in the lesson to illustrate the key ideas.
4. **Themes and areas of knowledge**. The TOK syllabus has five areas of knowledge and several "themes". Our book is organized thematically, which means that we don't discuss these elements one by one - instead, we discuss them all in comparison. However, if you want to understand how each lesson links to these elements of the TOK syllabus, it is stated here.
5. **Recap and plan.** It's a small section at the start of each lesson to give you a brief overview of what was discussed previously. It also introduces what will be discussed in the lesson.
6. **Boxes in the margins**. Each of these boxes contains a knowledge question that is related to one of the four elements of the IB knowledge framework (Scope, Methods and tools, Perspectives, Ethics). Sometimes these questions are directly discussed in the text, sometimes they are more of a "stop and think" point to extend your thinking. By their very nature, they are always debatable questions. Your teacher will choose to discuss some of these questions in class, leaving others for you to reflect upon on your own.
7. **Critical thinking extension**. This box at the end of the lesson is designed for students who are willing to explore more abstract problems of knowledge to exercise their critical thinking on a deeper level.
8. **If you are interested**. This box gives you suggestions regarding further reading or watching.
9. **Take-away messages**. This box at the very end of this lesson summarizes, just in one paragraph, the main ideas discussed in the lesson. It's a gist of the whole thing.

Units

The lessons are organized into larger units:

- **Introduction**. It contains three lessons explaining what TOK is and covering all essential curriculum terminology.
- **Unit 1: Knowledge of knowledge**. This unit is about knowledge itself - what is it, can it be defined, how are knowledge questions different from questions about the world?
- **Unit 2: Knowledge and technology**. This unit deals with the changing nature of knowledge in the age of technology. Can technology create a revolution in knowledge, change it beyond recognition? We discuss these questions in relation to all five areas of knowledge.
- **Unit 3: Bias in personal knowledge**. This unit explores one of the key concepts in the entire course - bias. Here we look at how bias influences knowledge in your everyday life. How do you know if you are biased or not, and is it possible for you to become less biased?
- **Unit 4: Bias in shared knowledge**. This unit continues exploring the concept of bias, but this time it is applied to three major areas of knowledge - Natural Sciences, History and Mathematics.
- **Unit 5: Knowledge and understanding**. This unit introduces such key concepts as objectivity and subjectivity, interpretation and understanding. What does it mean to understand something and how is understanding different from knowing? We apply this to Natural Sciences, Human Sciences and the Arts.
- **Unit 6: Knowledge and language**. This unit explores the role that language plays in both thinking and communication. Does language shape what we can know? Can we think without using a language? We also apply these problems to all five areas of knowledge.
- **Unit 7: Assessment guidance**. This unit contains focused advice on how to approach the TOK exhibition and TOK essay. We look at assessment instruments, analyze common mistakes and discuss checklists designed to ensure that you maximize your chances of getting the perfect marks.
- **Glossary**. This section contains an explanation of each of the key concepts used in the book.

Additional comments

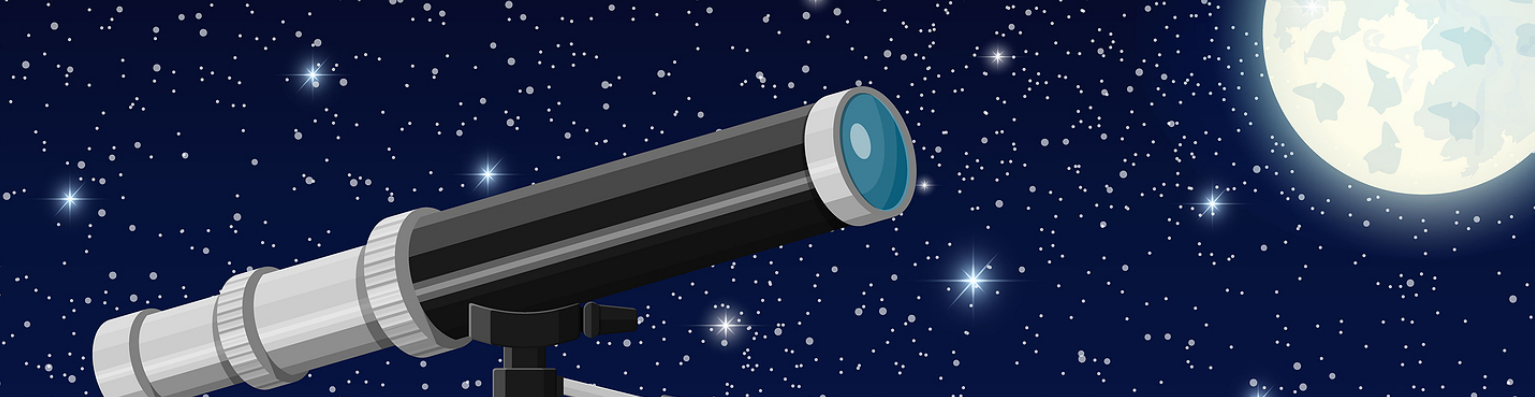
You will notice that each lesson, including all extension boxes, is maximum 1,600 words long. This is symbolic because that's exactly the word limit for the TOK essay.

Throughout the book I am modelling the kind of thinking that will be required of you in the assessment components. I ask questions and attempt answering them. You don't have to - and you shouldn't - agree with me on the conclusions I'm reaching. But it's the process of thinking that matters, the journey that took me there. Similarly, in your TOK essay and the exhibition it is not the conclusions that are assessed, but the process of thinking that you have demonstrated.

In each unit you will find one or more "Exhibition" and one or more "Story". These serve to demonstrate links between TOK and the real world.

Finally, you don't have to use this book sequentially. Each unit is relatively independent of other units, and each lesson is relatively independent of other lessons. This book is designed to be used in class, but it is equally suitable to be used at home when you are working on the arguments for your TOK assessments.

Enjoy!



INTRODUCTION

Lesson 1 - What is TOK? 8

Lesson 2 - Elements of TOK 12

Lesson 3 - Knowledge framework 18

Lesson 1 - What is TOK?

Learning outcomes

a) [Knowledge and comprehension] What is TOK about?
b) [Understanding and application] Why is it important to learn TOK?
c) [Thinking in the abstract] Why do we need meta-knowledge over and above regular knowledge?

Key concepts

Theory of Knowledge

Other concepts used

Epistemology, meta-knowledge

Themes and areas of knowledge

Theme: Knowledge and the knower
AOK: Mathematics, History, Natural Sciences

What is TOK about?

What you learn in various subjects at school is knowledge. For example, you learned in an Economics class that scarcity drives people to make decisions about how to allocate resources efficiently. That's your knowledge.

TOK is knowledge of knowledge. The main question that it attempts to answer is "How do we know what we know?"

For the example above, how do you know that it is scarcity that motivates people to allocate resources, not something else? How do you even know what drives people? It's not like you can see inside their minds. How reliable are statements like this? More generally, how universal are laws of economics? Are they more or less universal than laws of physics or chemistry? Is knowledge in economics as certain as knowledge in mathematics and, if not, why can't it be?

These and other questions would be examples of the things TOK explores. Unlike all other subjects where you gain knowledge about the world, in TOK you gain knowledge about knowledge about the world.

KEY IDEA: TOK is knowledge of knowledge. The main question that it attempts to answer is "How do we know what we know?"

Why learn TOK?

When I went to high school, my curriculum was very different from yours. I had 18 compulsory subjects. The way they were taught was less detailed than what you get in the IB, so I got broader coverage but less depth. Psychology and TOK, the two areas I ended up specializing in, were not part of my school curriculum. You might say that I have studied 18 different subjects just to discard them and pursue something else.

***Image 1**. Knowledge*

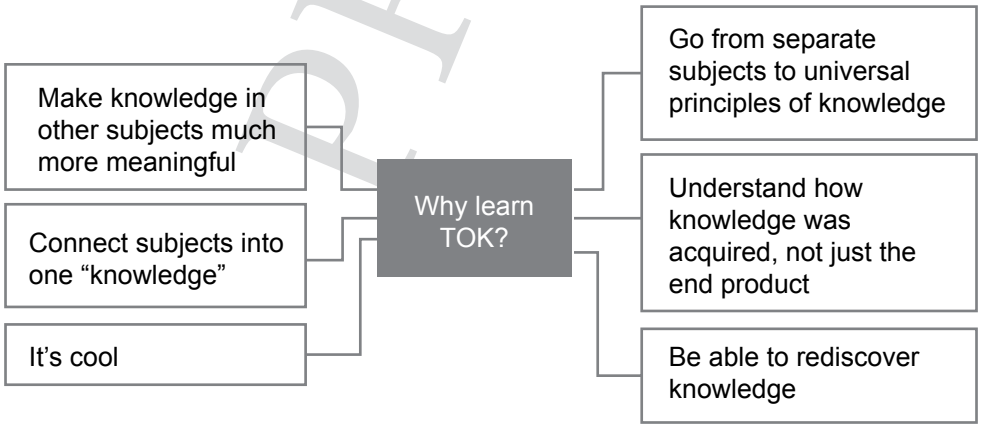
I think that as I was learning my 18 compulsory subjects, I always felt like I was lacking something. Perhaps it was some common understanding that would bring these subjects together, or some universal principles of knowledge. Back then I couldn't really put a label to it, but now I know – I was lacking a TOK course.

I learned about the Pythagorean theorem (Math), the Napoleonic Wars (History), Newtonian laws of motion (Physics). But I couldn't help asking myself "why?" and "how do we know?"; my education was not too helpful in providing these answers.

For the Pythagorean theorem, we were required to formulate it and be able to apply it in solving problems. It was fine. But I remember accidentally coming across a book that explained how the Pythagorean theorem was derived from the simple starting axioms. The proof was not difficult, and I was able to close the book and reproduce it on a sheet of paper. That moment changed my perception of mathematics. I realized I don't have to memorize the theorem; if I happen to forget it, I can simply reconstruct the proof. Now that I knew where my knowledge came from, it felt so much deeper. Do you have your own examples of when you learned how a particular piece of knowledge was discovered, after which this knowledge suddenly made much more sense to you?

For the Napoleonic Wars, I was told what happened, when and how. I was given the end result of the work of a historian, but I was never required to play the role of a historian myself. Years later, I had to find out what happened to IQ testing in the 1930s in Soviet Russia and why it was banned for decades. I looked at the heap of documents that I managed to find and wondered how a historian can ever make sense of all this. Have you ever tried writing history? Try writing down how the current leader of your country came to power in one paragraph, and you will understand the tremendous amount of mental work that goes into this paragraph.

For Newtonian rules of motion, I was given the formulas and expected to take them for granted. I learned later that Newtonian laws are based on one important assumption: that the body is moving in an "inertial space" where no other forces exist. But I also learned that in real life, inertial space doesn't exist. So does it mean that his equations do not fully apply to the real world? More importantly, what other knowledge from Physics did I take for granted without questioning the assumptions upon which it is based? Do an exercise: remember one piece of knowledge that you studied in Physics (either in the IB Diploma Programme or before that) and identify an assumption upon which this knowledge is based. How easy is that for you?



But over and above this reflection on the limitations of knowledge that I was getting in my 18 separate school subjects, I lacked something that would meaningfully combine these subjects into one "knowledge". After all, academic disciplines are divided into subjects, but the real world is not. When you are reading this, your neurons are firing electricity (physics), your

brain is producing chemical messengers (chemistry), your heart pumps blood to send oxygen to the parts of your brain that are active (biology), and you engage in the mental process of reading and understanding using language (psychology). This is one single process, but we break it down and study its components separately in separate subjects.

When I discovered Theory of Knowledge, it made my knowledge in all other subjects much more meaningful.

What is TOK like?

Theory of Knowledge is a special subject. It has critical thinking written all over it. Depending on how you approach the subject, it may either leave you with a puzzling aftertaste ("What was that???") or entirely change the way you think ("That is so cool, I'm going to do it all the time!"). Obviously, we want to achieve the latter. However, being thoroughly puzzled about something is a necessary part of changing the way you think. If you do not feel puzzled or perplexed, you are not really challenging what you already know. Hence, you are pursuing an illusion of knowledge, but not knowledge itself. Therefore, I encourage you to be confused as often as you possibly can.

The first humans dramatically advanced in their development when they started using tools. Cooking food was easier with fire, hunting was easier with a spear and transportation was easier with the wheel. Tools allow us to explore the reality of the physical world. In a similar way, there are tools that help us explore the reality of the mental world (the world of knowledge). These tools are concepts. We use concepts to think about the world and ourselves, and concepts become lenses through which we know. The cleaner the lenses, the more clearly we understand things.

This is why this course is conceptual. It is designed around such central concepts as doubt, justification, truth, evidence, and so on. If you clearly understand these concepts, you will be able to apply them to various domains of knowledge and understand these domains better than ever before. There is no memorization involved in the course, but a lot of questioning, understanding and application.

***Image 2**. Knowledge is power*

Critical thinking extension

The prefix "meta" has roots in ancient Greek where it meant "after" or "beyond". You might recall a lot of instances where you have come across "meta"-something. Here are several examples:

- Metacognition in psychology means cognition about cognition (for example, when you think about how you can remember exam material better).
- Metadata in computer jargon means data about data (for example, data for Twitter is the text of the tweets while metadata is information on when and where the tweet was posted).
- Metaphysics is sometimes used synonymously with "philosophy". Aristotle originally divided disciplines into Physics (the study of nature) and metaphysics (after the Physics).

Theory of Knowledge deals with "meta" a lot. If your other school subjects are all about knowledge, then TOK is all about meta-knowledge. To what extent do you think a "meta"-something is necessary to fully understand this something? Can you come up with examples?

If you are interested…

Another term for theory of knowledge is "epistemology". In fact, this is exactly how the word "epistemology" is translated from its ancient Greek roots: *epistēmē* = knowledge, *logos* = study or theory.

All philosophy may be very broadly divided into two parts:

- Ontology (theory of being). This focuses on claims such as "X is" or "X exists". For example, God exists, infinity exists, the Universe is infinite.
- Epistemology (theory of knowledge). This focuses on questions such as "How do we know X is?" or "How do we know X exists?" For example, "How do we know that the Universe is infinite?"

IB TOK is not philosophy, though. We are staying away from all technicalities and nuances of philosophy and instead we are focusing on applications of knowledge concepts to specific areas of knowledge.

However, there is certainly a lot of overlap between IB TOK and epistemology as a branch of philosophy.

Take-away messages

Lesson 1. The main question that we attempt to answer in TOK is "How do we know what we know?" TOK is a reflection on our knowledge, a knowledge of knowledge. The value of TOK may be seen in understanding the deep underlying principles that govern the acquisition of knowledge in various areas, such as human sciences, mathematics, the arts. Additionally, TOK allows us to have a basis upon which various disciplines can be compared and combined. The division of knowledge into academic disciplines is artificial (it does not exist in the real world), and TOK is trying to restore the balance by tying them all back together. TOK is a conceptual subject. At its core are conceptual understanding and critical thinking.

Lesson 2 - Elements of TOK

Learning outcomes

a) [Knowledge and comprehension] What key elements does the course consist of?
b) [Understanding and application] What is the role of themes in the course?
c) [Thinking in the abstract] How can we draw a line between personal knowledge and shared knowledge?

Key concepts

The knower, personal knowledge and shared knowledge, areas of knowledge, knowledge questions, knowledge framework, themes

Themes and areas of knowledge

Themes: Knowledge and the knower, Knowledge and language, Knowledge and technology
AOK: Natural Sciences, Human Sciences, Mathematics, History, the Arts

Recap and plan

We have discussed what TOK is, what it "feels like" and why it is important to learn it at school. Now we will have an overview of the main components of the IB TOK course.

The knower

In the center of TOK is the knower – a person who knows. I am a knower, you are a knower. But we also belong to various communities of knowers, such as the community of people sharing a particular religious belief, the community of mathematicians, the community of students who learn European history from European textbooks.

Personal knowledge and shared knowledge

The knower has certain knowledge about himself/herself and the world around them. This knowledge can be of two types:

- Personal knowledge
- Shared knowledge

Personal knowledge is something that belongs to an individual and is not necessarily shared by other individuals. Shared knowledge is something that is jointly produced by large groups of people. Such knowledge is common to large communities. For example, mathematics is in the domain of shared knowledge. On the other hand, your intuitions about different types of food and how tasty they are belong to the domain of your personal knowledge. It may or may not be shared by others. Similarly, physics is shared knowledge, but a student's understanding of physics is that student's personal knowledge.

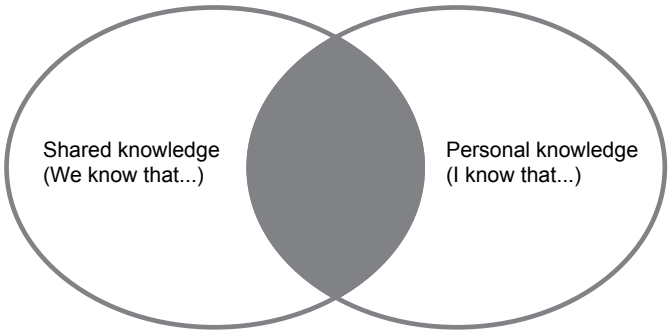


***Image 3**. Personal and shared knowledge*

Areas of knowledge

Shared knowledge may be further divided into areas of knowledge (AOKs). In IB TOK, we speak about five such areas:

- Natural Sciences
- Human Sciences
- Mathematics
- History
- The Arts

These areas of knowledge may be distinctly different in many aspects. Comparisons between these areas of knowledge through a conceptual lens is what comprises the bulk of the IB TOK course.

Knowledge questions and knowledge claims

The main focus of the course is on *knowledge questions* and *knowledge claims*.

Knowledge questions are questions about knowledge itself, such as "What counts as good evidence for a claim?" or "Are some types of justification more reliable than others?" Since these are questions about knowledge itself, they draw on TOK concepts rather than subject-specific terminology. Knowledge questions are contestable, in the sense that the answer to them is not obvious and there may exist various reasonable approaches to an answer.

A knowledge claim is a statement in response to a knowledge question. For example, "The quality of evidence is determined by its consistency with previous knowledge" or "Justifications based on observation are more reliable than logical proofs".

Knowledge framework

In IB TOK, knowledge questions are broadly organized into four categories. You may think of them as "groups" of knowledge questions. The categories, known as the knowledge framework, are:

1) Scope
2) Perspectives
3) Methods and tools
4) Ethics

***Image 4**. Knowledge framework*

It is a requirement of the course that all four groups of knowledge questions are discussed.

You should not worry too much about which question belongs to which category. Sometimes categories overlap and one knowledge question may belong to more than one category. You are not required to "correctly" place knowledge questions under categories, but you are required to ensure that all four categories have been discussed. This way the IB makes sure that you do not skip, say, ethics.

In the next lesson we will discuss in more detail the nature of each of these four elements, as well as their applications in the five areas of knowledge.

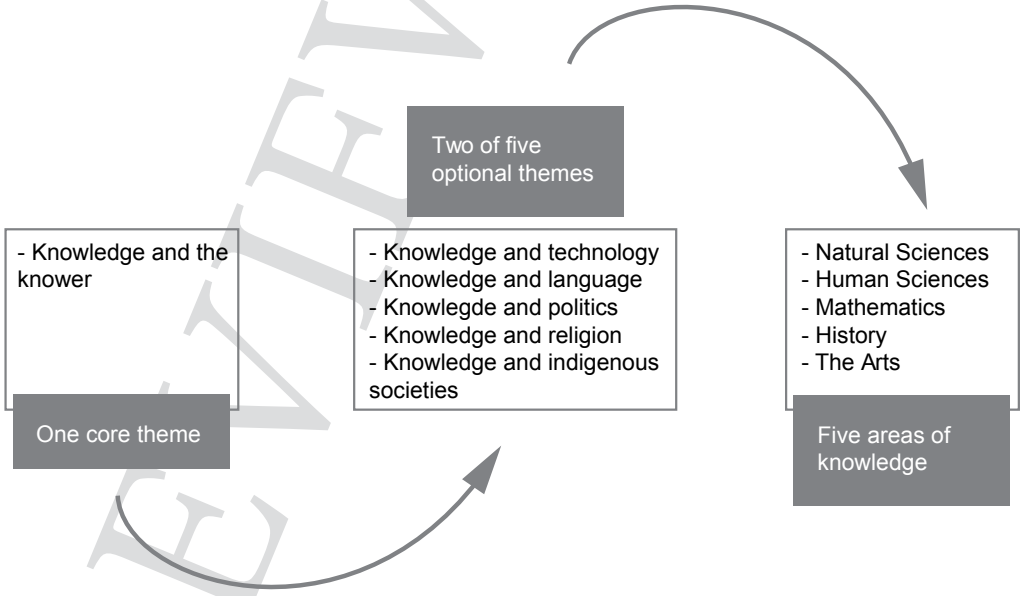
Themes

Apart from the five areas of knowledge, students in IB TOK are required to study three themes: the core theme and two of five optional themes.

The core theme is "Knowledge and the knower". It is focused on personal knowledge. It is a reflection on yourself as a knower and thinker.

The five optional themes are:

1) Knowledge and technology
2) Knowledge and language
3) Knowledge and politics
4) Knowledge and religion
5) Knowledge and indigenous societies



How the Themantic course is organized

Themantic Education designs courses with a focus on conceptual understanding and continuity of knowledge. We do not like the idea of studying each area of knowledge separately, one after another. Instead, we are looking at key TOK concepts and discussing how they manifest in various areas of knowledge. This allows for effective comparisons.

This book is organized around our own broad "themes". Here is a brief summary of our themes and how they map onto the elements of IB TOK:

Our themes	IB guide themes	Natural Sciences	Human Sciences	Mathematics	History	The Arts
Introduction.		✓	✓	✓	✓	✓
Unit 1. Knowledge of knowledge		✓		✓	✓	
Unit 2. Knowledge and technology	Knowledge and technology	✓	✓	✓	✓	✓
Unit 3. Bias in personal knowledge	Knowledge and the knower		✓			
Unit 4. Bias in shared knowledge		✓		✓	✓	
Unit 5. Knowledge and understanding		✓	✓			✓
Unit 6. Knowledge and language	Knowledge and language	✓	✓	✓	✓	✓
Unit 7. Assessment						

In this book we are discussing areas of knowledge not after themes and not separately from them, but *through* themes.

Extra themes that we added (knowledge of knowledge, knowledge and understanding, bias) make it possible to compare areas of knowledge conceptually within a meaningful framework. Each theme will be organized around important concepts that have relevance to all areas of knowledge. This will allow us to compare areas of knowledge throughout the book.

Assessment

In TOK there are two assessment components: a TOK exhibition (internal assessment) and a TOK essay (external assessment).

***Image 5**. Assessment*

For the exhibition, you explore how TOK manifests in the world around us. There are 35 IA prompts (formulated as knowledge questions). You are required to select one of the 35 prompts and center your exhibition around this prompt. Examples of IA prompts are:

- (IA prompt 12) Is bias inevitable in the production of knowledge?
- (IA prompt 19) What counts as a good justification for a claim?
- (IA prompt 32) What makes a good explanation?

You will find the full list of prompts in the IB TOK Guide.

Your exhibition should comprise of three objects (or images of objects) plus a written commentary on each object (a maximum of 950 words for all three commentaries combined). In the commentary, you are required to identify the object and explain its real-world context and its connection to the IA prompt.

The exhibition is internally assessed and externally moderated. It is worth 35% of your marks.

For the essay, six months prior to the submission deadline the IB releases "prescribed essay titles". You are required to choose one of these titles and write an individual essay on it (word count limit is 1600 words). This is an external component marked by IB examiners. It is responsible for 65% of your marks. The essay title will be formulated as a knowledge question. You are assessed on the quality of your argumentation, consideration of different points of view, and making links to areas of knowledge.

You can find further guidance on TOK assessment in Unit 7 of this book.

Perhaps the most important thing that you need to understand at this point is that TOK is not assessed in a conventional way. There is nothing to memorize. It is all about understanding and thinking. It is also about skills. It is impossible, for example, to predict what the prescribed essay title will be, so it is highly likely that you will have to write an essay on something that you never discussed in class. That being said, you have plenty of time to do so and will be able to do your research if necessary. But knowledge of key concepts will help you immensely.

You should use this book accordingly. Understand the concepts, do the thinking, argue and disagree. Content only matters as far as it enables good argumentation, and your knowledge of content itself will not be assessed.

Critical thinking extension

There is a complex relationship between personal knowledge and shared knowledge. The boundary between these two is not always clear.

In fact, the IB does not officially use the terms "shared knowledge" and "personal knowledge" in the Guide (they used to be there in the previous syllabus), but this distinction is implied. Areas of knowledge are about "shared knowledge". The core theme is about personal knowledge. The optional themes may cover both aspects.

For our course it is useful to return to the clear distinction between personal and shared knowledge. We will be alternating between them from time to time, and it is important that you bear in mind the profound difference between "I know that…" and "We know that…".

Can you think of several things that you know that are uniquely your own, several things that you know differently from your classmates, several things that you know because you belong to a certain knowledge community? Where do you think we should draw a line between personal knowledge and shared knowledge?

If you are interested…

The IB TOK Subject Guide is the official IB publication outlining the syllabus, assessment requirements and other important details of the course. While your teacher, just like any other IB TOK teacher in the world, follows the Guide closely, it may be a good idea for you to also familiarize yourself with this document and have ready access to it. Ask your teacher to share it with you.

Take-away messages

Lesson 2. The key components of the TOK course are the knower, personal and shared knowledge, knowledge questions, knowledge framework, areas of knowledge, and themes. Rather than looking at each area of knowledge separately, this book looks at areas of knowledge through themes. This allows us to compare and combine areas of knowledge within the key concepts. Assessment in the course includes two components: the TOK exhibition (internal assessment) and the TOK essay (external assessment).

Lesson 3 - Knowledge framework

Learning outcomes

a) [Knowledge and comprehension] What elements does the knowledge framework consist of?
b) [Understanding and application] What role does the knowledge framework play in the TOK course?
c) [Thinking in the abstract] How should we treat knowledge questions that can be related to more than one category?

Key concepts

Scope, methods and tools, perspectives, ethics

Themes and areas of knowledge

Themes: Knowledge and the knower, Knowledge and technology, Knowledge and language
AOK: Natural Sciences, Human Sciences, History, Mathematics, the Arts

Recap and plan

In the previous lesson we looked at the main components of the IB TOK course. One of these elements – the knowledge framework – requires a closer look.

As you already know, the course revolves around knowledge questions, and knowledge questions may be broadly organized into four groups: scope, perspectives, methods and tools, ethics.

So what is the focus of each of these four elements of the knowledge framework?

Scope

This element explores the nature of the problems that are investigated in each theme / area of knowledge. It also shows the place of the theme / area of knowledge within human knowledge in general.

***Image 6**. Scope*

Examples of questions relating to scope are:

- What are the key unanswered questions and unsolved problems currently in this area of knowledge (or theme)?
- What makes this theme or area of knowledge important?

There will be more specific knowledge questions related to scope within each theme and area of knowledge. Examples are given in the table below:

Theme / area of knowledge		**Examples of knowledge questions related to scope**
Themes	Core theme: Knowledge and the knower	Is there a limit to how far we can know ourselves? How biased is our personal knowledge?
	Knowledge and language	Is it possible to have knowledge without language? Can all knowledge be expressed in language?
	Knowledge and technology	How has the development of technology influenced the way we know things? Can computers make discoveries on their own?
Areas of knowledge	Natural Sciences	Is there anything that is beyond scientific understanding? What counts as scientific knowledge?
	Human Sciences	Can human sciences be replaced by natural sciences? What is it about humans that makes them a special object of research as compared to other areas of knowledge?
	Mathematics	Is mathematics a study of abstract entities or a study of the real world? How does technology affect the nature of mathematical knowledge?
	History	Is there a difference between knowledge and interpretation in history? Is knowledge of the past useful for the present?
	The Arts	What counts as knowledge in the arts? Is the aesthetic value of an artwork universal or a matter of personal opinion?

Perspectives

This element of the knowledge framework focuses on the possibility of varying interpretations or points of view regarding knowledge of something. When knowledge is open to interpretation and there are several ways of looking at it, perspectives come into play.

The table below gives some examples of knowledge questions related to this element of the knowledge framework in various themes and areas of knowledge:

***Image 7**. Perspectives (credit: Mushki Brichta, Wikimedia Commons)*

Theme / area of knowledge		**Examples of knowledge questions related to perspectives**
Themes	Core theme: Knowledge and the knower	Are personal beliefs determined by personal experiences? Is it inevitable that my knowledge will always be biased in one way or another?
	Knowledge and language	Does language contain knowledge or does it merely express it? Are there universal concepts shared by humans which are not likely to be shared by aliens?
	Knowledge and technology	Is human knowledge fundamentally different from products of computer algorithms? Does modern technology create paradigm shifts in areas of knowledge?
Areas of knowledge	Natural Sciences	Does scientific progress get us closer to the truth? Is there such a thing as an objective scientific fact?
	Human Sciences	Is it possible to understand subjective human experiences objectively? Is bias in human sciences desirable in any way?
	Mathematics	How is mathematical knowledge related to the real world? Can mathematics be biased?
	History	Is a historical perspective the same as bias? Does a combination of perspectives allow us to get closer to a historical truth?
	The Arts	Can knowledge conveyed by a work of art be universal to all people? Is art knowable? Does technology change the nature of art?

Methods and tools

This element explores how knowledge is produced. Different areas of knowledge as well as individual knowers can use different ways of obtaining knowledge. This is not limited to formal methodologies (for example, the experimental method or the deductive proof), but also includes cognitive tools (such as assumptions, analogies, reasoning, perception). Technology can also serve as a tool for producing knowledge.

***Image 8**. Tools (credit: Styx, Wikimedia Commons)*

Examples of knowledge questions related to this element of the knowledge framework can be found in the table below:

Theme / area of knowledge		**Examples of knowledge questions related to methods and tools**
Themes	Core theme: Knowledge and the knower	How do we acquire knowledge about ourselves and the world around us? How can we overcome our own bias?
	Knowledge and language	How does language make it possible to manipulate beliefs and opinions? Can we think beyond concepts that we have internalized together with language?
	Knowledge and technology	How does technology overcome limitations of human knowledge? Are there aspects of the world that can be understood only by using computer simulations?
Areas of knowledge	Natural Sciences	How important is it to establish causation in scientific knowledge? Can we accept claims in natural sciences if they cannot in principle be confirmed by observation?
	Human Sciences	What does it mean to "understand" in human sciences, as compared to other areas of knowledge? Is the use of subjective methods in human sciences justifiable?
	Mathematics	How does constructing axiomatic systems differ from constructing scientific knowledge? Can computers prove theorems?
	History	How can we go beyond reporting events of the past to reconstructing their meaning? Does Big Data provide a fundamentally different approach to constructing historical knowledge?
	The Arts	How important is it to know the context to understand a work of art? What is the essential difference between knowledge of art critics and that of laypersons?

Ethics

This element of the knowledge framework explores knowledge questions implied in the ethical issues that arise in the process of obtaining knowledge. Note that the focus is not on the ethical issues themselves, but on the wider understanding of the relationship between knowledge and ethics.

***Image 9**. Ethics*

Some more specific examples from themes and areas of knowledge are given in the table below:

Theme / area of knowledge		**Examples of knowledge questions related to ethics**
Themes	Core theme: Knowledge and the knower	If there is a bias in our knowledge we are not aware of, do we still bear moral responsibility for negative consequences of this bias? Are we obligated to share what we know?
	Knowledge and language	Who is responsible for misunderstanding occurring as a result of using language? How can we know when language is misused for purposes of manipulation?
	Knowledge and technology	Is data privacy more important than knowledge that could be gained if all data were open? Is it our moral obligation to try and develop artificial consciousness because it can allow us to understand ourselves?
Areas of knowledge	Natural Sciences	Should ethical considerations constrain scientific research? Can natural sciences explain morality?
	Human Sciences	Should human sciences be descriptive or prescriptive? In what ways can ethical considerations be said to enhance knowledge in human sciences?
	Mathematics	Are ethical principles similar to mathematical statements that logically follow from a set of assumptions? What are the ethical issues surrounding commercial licensing of software used to prove theorems?
	History	Is it fair to apply modern standards to judge people of the past? Do historians have a moral responsibility to eliminate their own perspectives from their account of the past?
	The Arts	Are aesthetic judgments similar to ethical judgments? Are there any circumstances in which the unethical can be beautiful?

Critical thinking extension

Overlap between elements

You must have noticed that sometimes there is considerable overlap between elements of the knowledge framework. At times, one and the same knowledge question could be reasonably related to more than one category. For example, the question within natural sciences asking "Is there such a thing as an objective scientific fact?", depending on the angle at which we look at it, may be related to:

- Scope (whether or not "objective facts" lie within the scope of natural sciences)
- Perspectives (there are arguments for and against)
- Methods and tools (because we use the scientific method to be able to claim that something is an "objective fact")

What other examples can you identify?

Don't worry about these overlaps. They are perfectly natural because in real life, knowledge is not broken down into artificial categories. You will not be assessed on how "correctly" you can place various knowledge questions in various categories.

If you are interested…

There are many more examples of knowledge questions in the IB TOK Subject Guide. You might want to take a look at them, especially focusing on the themes and areas of knowledge you are more familiar and comfortable with, to get an idea of the range and type of knowledge questions that could be explored in the course.

Take-away messages

Lesson 3. The knowledge framework is a tool IB TOK uses to group knowledge questions into categories. There are four such categories: scope, perspectives, methods and tools, ethics. The knowledge framework is meant to ensure that for each area of knowledge and each theme, students discuss knowledge questions related to all four categories. This prevents a one-sided exploration of areas of knowledge. It is not always easy to place a knowledge question under one of the four categories unambiguously, but this is not what is required. In this lesson, we looked at examples of knowledge questions for each category in each of the themes and areas of knowledge.

UNIT 2 - Knowledge and technology

Exhibition: Graph of emotions in the Bible 73
Story: Predicting Supreme Court decisions 74

2.1 - Technology and personal knowledge 76
Lesson 1 - Information bubble 76

2.2 - Technology and the human mind 80
Lesson 2 - AI: Turing test 80
Lesson 3 - AI: Artificial consciousness 85
Lesson 4 - Hard problem of consciousness 90
Lesson 5 - Technological singularity 94

2.3 - Technology in Natural Sciences 99
Lesson 6 - Computer simulation 99
Lesson 7 - Simulated world 104
Lesson 8 - Computer-generated knowledge 109

2.4 - Technology in Human Sciences and History 113
Lesson 9 - Big Data 113
Lesson 10 - Nomothetic and idiographic research 118
Lesson 11 - Text mining 123

2.5 - Technology in Mathematics 127
Lesson 12 - Proof-by-exhaustion 128
Lesson 13 - Experimental mathematics 133

2.6 - Technology in the Arts 137
Lesson 14 - Redefinition of art 138
Lesson 15 - Digital art 142

2.7 - Technology and ethics 146
Lesson 16 - Technoethics 146

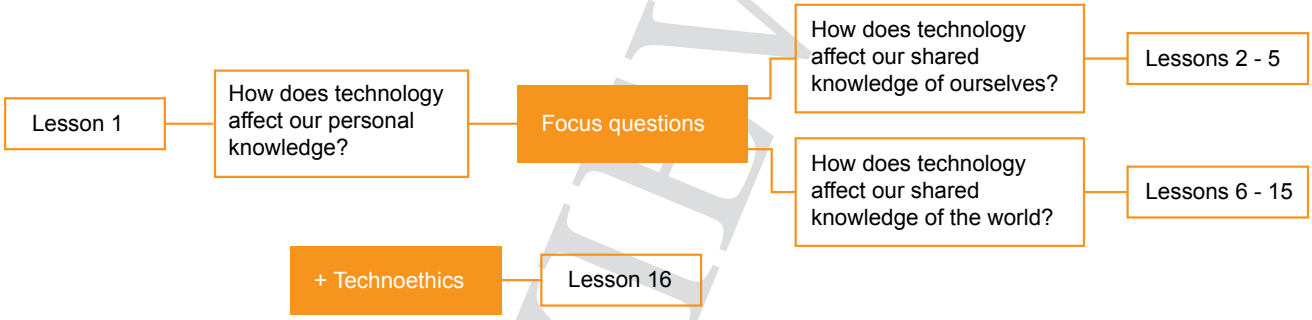
Back to the exhibition 151

UNIT 2 - Knowledge and technology

Technology has become such an integral part of our lives, and is changing our lives so deeply, that I'm having a hard time choosing the areas of focus for this unit. There is so much to explore.

After giving it some thought, I decided to concentrate on the following questions:

1) How does technology affect our personal knowledge? Now that information is so readily and instantly accessible to us, does it change how we know things?
2) How does technology affect our shared knowledge of ourselves?
3) How does technology affect our shared knowledge of the world?



The first question (lesson 1) is about how technology has transformed the way you and I go about acquiring knowledge in our everyday lives. Admittedly, we have much easier access to information now that we have the Internet. But this is only the tip of the iceberg. Technology may be influencing our knowledge acquisition in negative ways, too. For example, search engines these days are proactive: they return results that they "think" we will find interesting. Therefore, they make some important decisions regarding relevance of information for us. We have outsourced these decisions to them. This may negatively affect us because we are trapping ourselves in an information bubble.

Note that the first question is about personal knowledge. New technology poses numerous challenges to every individual knower, but I assume that collectively we can overcome these challenges (although it may be difficult). The second and the third questions are questions about shared knowledge.

The second question (lessons 2-5) is about how technology invented by human beings has allowed human beings to understand the phenomenon of human beings. Our own brains, minds and consciousness are perhaps the toughest puzzle of the Universe. We have many questions in this area that we cannot even begin to approach answering. But if we manage to build a machine that can think, act and perhaps even feel like a human being, then we can claim to have understood these phenomena. This question revolves around artificial intelligence: what it looks like today, what it will look like in the future and how the relationship between humans and machines is likely to transform within our lifetime.

The third question (lessons 6-15) is about how we (collectively, as humanity) can use technology to better understand the world around us. There are simple and obvious examples that come to mind in response to this question: we invented the microscope and were able to see the living cell; we invented the telescope and were able to see distant galaxies; we invented brain imaging and were able to see inside the living brain without cutting the skull open. But there are also ways in which technology might have changed our knowledge of the world more radically. Does technology have the potential to create revolutionary changes in the areas of

Natural Sciences, Human Sciences, History, Mathematics and the Arts? This question is about the numerous tools we have invented to acquire new knowledge about the world as well as the strengths and limitations of these tools.

Another aspect that weaves through all of these questions is ethics (mentioned throughout the unit, but revisited in lesson 16). Technology has enabled us to do things we were not able to do before. But just because we *can* do something, does it mean that we *should* do it? Genetic engineering, for example, allows us to modify or edit genes in order to change an organism's characteristics in a particular way. We can (theoretically) clone people. We can even train an artificial intelligence to write TOK essays and submit them as our own. But should we do all that? Technology bears certain dangers. Is it anyone's responsibility to assess risks and prevent disasters?

Exhibition: Graph of emotions in the Bible

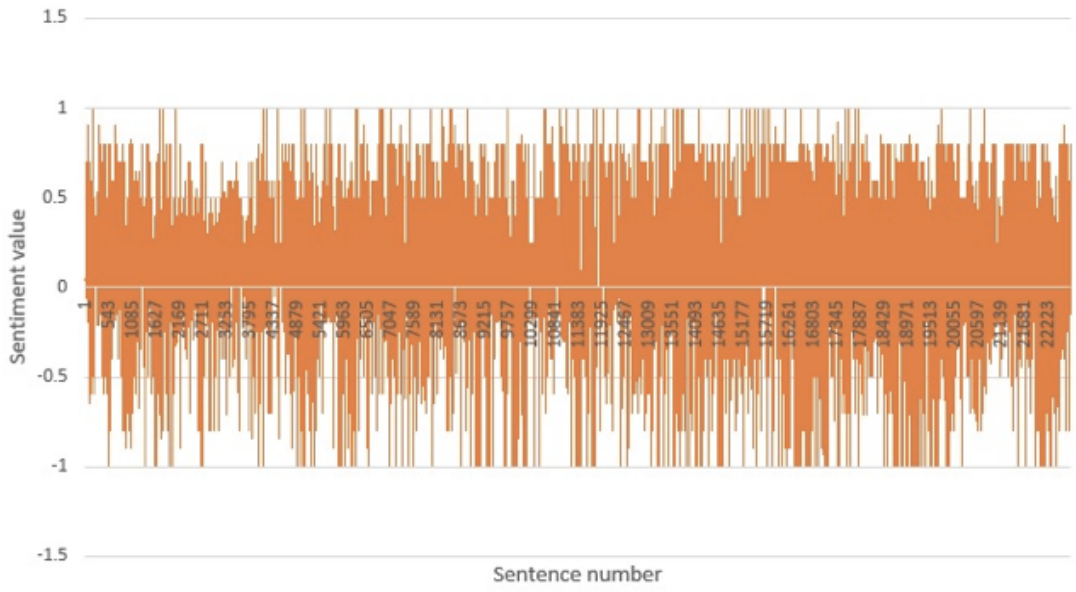


***Image 1**. Bar graph of emotions in the Old Testament (sentiment analysis)*

This is a graph of emotions in the Bible.

At some point in my career I spent several years working as a data scientist for a large company. I was trying to get an insight into the behavior of people by crunching numbers. One of the things that sparked my interest was the method of text mining known as "*sentiment analysis*".

This is how it works:

1) A group of researchers takes a large collection of words (nouns, verbs, adjectives) from a dictionary.
2) Then they ask a group of participants to rate the emotional valence of each word. Participants rate each word on a continuous scale from -1 (very negative) to +1 (very positive). For example, words like "sun", "pleasant" and "hugging" will get high positive ratings, while words like "mutilation", "dirty" and "vandalize" will get high negative ratings. This results in a database of words and their average emotional ratings.
3) If you want to calculate sentiment value of a text, you run this text through the database. The score for the text will be calculated based on the concentration of "emotional" words in it. For example, if the text contains a large number of positively colored words, then it will score a high positive sentiment value.
4) You can run multiple texts through sentiment analysis and compare sentiment values of these texts.

There are multiple examples of how sentiment analysis is used in both business and research. One such example is sentiment analysis of tweets from politicians. Every tweet is a text, so we can run sentiment analysis and calculate the sentiment value for each tweet. But every tweet also contains some useful metadata: the timestamp (when it was posted), a geotag (where it was posted from), what device it was posted from, and so on. You can then play around with data. For example, we could visualize the amount of negativity in a president's tweets depending on where they are travelling.

In my spare time, I used sentiment analysis for a more modest purpose – to analyze sentiment of the Bible. I downloaded a copy of the Old Testament from the Internet (easy to find!). I installed Python (the programming language). Then I broke down the Old Testament into separate sentences, so that each sentence becomes one unit of text. I then ran a sentiment analysis on each sentence and graphed the result. The whole thing took me just a dozen lines of code, by the way, and it was easy to do because there are plenty of step-by-step instructions online.

So, my graph of emotions in the Bible shows how the sentiment of sentences develops from the beginning to the end of the Old Testament. You can see that the Bible can get pretty positive sometimes; however, the happy notes don't reach too high and don't last for too long. On the other hand, when the Bible gets negative, it really does go all the way down. It reaches very low values of sentiment (close to -1) and stays there for a longer time.

Does my sentiment analysis of the Bible provide new insights and open up new horizons of knowledge that cannot be achieved by ordinary methods? If I put my graph in a frame and display it in an art gallery, will it deserve to be considered a proper work of art? Is it even ethical to treat a religious text as a dataset?

I seek your help in answering these questions.

Story: Predicting Supreme Court decisions

***Image 2**. U.S. Supreme Court building (credit: Wikipedia)*

This is a story of competition between legal experts and a computer algorithm. The focus of the competition was forecasting the U.S. Supreme Court decisions. The question was, are human law experts better than a simple computer program in predicting the outcome of cases heard in the Supreme Court?

Andrew Martin and Kevin Quinn analyzed data from 628 cases previously decided by the U.S. Supreme Court justices. For each of these cases, they collected only six simple observable characteristics, for example, the type of petitioner (the United States, an injured person, an employer, etc.), whether or not the petitioner appealed to the Constitution, where the case came from, and so on. There was no theory behind selecting these variables. They were selected simply due to their easy availability in public sources. To give you a sense of the rules that the algorithm operated with, here is one example: "If the petitioner was an injured person, if the petitioner did not appeal to the Constitution, and if the case came from the Federal Circuit, Justice Sandra Rey will vote to affirm".

After training the algorithm on prior data, they used it to forecast the outcomes of new cases. They then held a competition between their algorithm and human legal experts! All experts had extensive training and experience in their domain. There was a total of 83 experts, each an accomplished professional. Many of them had practiced or clerked at the Supreme Court. Experts were asked to forecast the outcome of the cases that were within their immediate area of expertise.

Martin and Quinn set up a public website where they placed their bets (voila, the website is here: http://wusct.wustl.edu/). On the website they announced the two sets of predictions (one from legal experts and one from their algorithm) *before* the hearing of the case. After the hearing, they recorded the outcome. They collected data in this manner for the duration of one year, the U.S. Supreme Court's 2002 term. So, were expert predictions better than the predictions of the simple algorithm manipulating six easily registered characteristics of each case?

No. The model was correct in 75% of the cases, while the experts were correct in 59% of the cases. The algorithm won convincingly and publicly.

Why?

One possible explanation is that human cognition is limited. When legal experts review a case and make a prediction, they base their forecast on prior experience. But when it comes to human beings, "experience" means a handful of cases that stand out for them, cases they can hold in their memory and process. The computer algorithm in this project was able to base the predictions on the total number of available cases – 628 – without any bias in selecting them, without giving them unreasonable subjective weight.

But even so, it is surprising the algorithm won because it did not take into account any legal explanations provided by the Court (and no legal interpretation at all, for that matter). The six variables were essentially non-legal. They did not consider any substance about the case, only "superficial" characteristics such as the type of respondent and where the case came from.

Ian Ayres, who describes this project among a dozen other examples in his book *Super Crunchers: Why Thinking-by-Numbers is the New Way to be Smart* (2007), comes to the conclusion that statistical algorithms are simply better at predicting than human experts.

So, should we sack all experts?

2.2 - Technology and the human mind

We have discussed the relationship between technology in personal knowledge and it is now time to switch over to shared knowledge.

In the next four lessons, we will be answering the question "How does technology affect our shared knowledge of ourselves?"

The key concepts here are artificial intelligence and artificial consciousness. It would be a mistake to think that this discussion is limited in its relevance to technology. It has much broader implications, including our understanding of what it means to be human and, ultimately, our answer to the question "Who are we?"

You see, if we manage to construct an artificial consciousness, this would mean that we have constructed a human being. This would mean that we fully understand what it means to be human. This would probably cause natural and human sciences to merge. This would irreversibly change the nature of knowledge in general.

Therefore, the questions dealt with in this part of the unit are not only related to understanding technology. They are equally related to our understanding of the human mind.

Lesson 2 - AI: Turing test

Learning outcomes

a) [Knowledge and comprehension] What is the Turing test?
b) [Understanding and application] Can machines act like they are intelligent?
c) [Thinking in the abstract] Should we strive to make artificial intelligence similar to human intelligence?

Key concepts

Artificial intelligence, Turing test, general intelligence, explicit and implicit thinking

Other concepts used

Chatbots, Loebner prize, personal assistants, thought experiment, brain, symbol manipulation

Themes and areas of knowledge

Theme: Knowledge and technology
AOK: Natural Sciences, Human Sciences, Mathematics

Recap and plan

We have looked at the role of technology in the ongoing transformation of our personal knowledge. In this new co-existence with our digital devices, they are becoming an extension of our brain.

But how far will this process go? Will we merge with machines? Will machines take over? Will we co-exist in a kind of knowledge symbiosis?

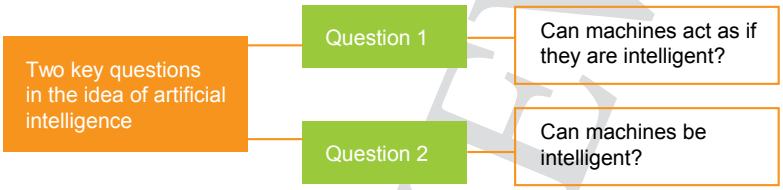
These are interesting and complicated questions, but to effectively address them we need to unpack such concepts as artificial intelligence and artificial consciousness over the next couple of lessons.

Two questions of artificial intelligence: acting intelligently and being intelligent

There are two key questions in the idea of artificial intelligence:

1) Can machines act as if they are intelligent?
2) Can machines be intelligent?

The difference between these two questions is really important. Mixing them up leads to a lot of confusion in any AI-related conversation. To keep them clearly separate, in this lesson we will only deal with the first question.



Turing test

The famous Turing test proposed in 1950 is probably the best-known thought experiment in this area. Imagine you are in one room and in two other rooms there is a computer and another human being. You communicate with them via questions and answers. You write your question on a card and push it through a slot in the wall. Sometime later, two cards with answers come back, one from the computer and one from the human (but you don't know which one is from whom). You can spend some time asking questions related to a certain subject area and receive answers in return, and then you are asked to decide which answers came from the machine and which answers came from the human. If you are unable to do so, the machine is said to have passed the Turing test for artificial intelligence. It has fooled you into believing that it's human, therefore it can act intelligently.

Actually, according to Alan Turing himself, it also means that the machine is intelligent. He did not see a difference between the two questions above ("Can machines act like they are intelligent?" and "Can machines be intelligent?"). According to him, the only way we can understand that someone (or something) is intelligent is if that someone (or something) acts intelligently.

So, do machines currently pass the Turing test?

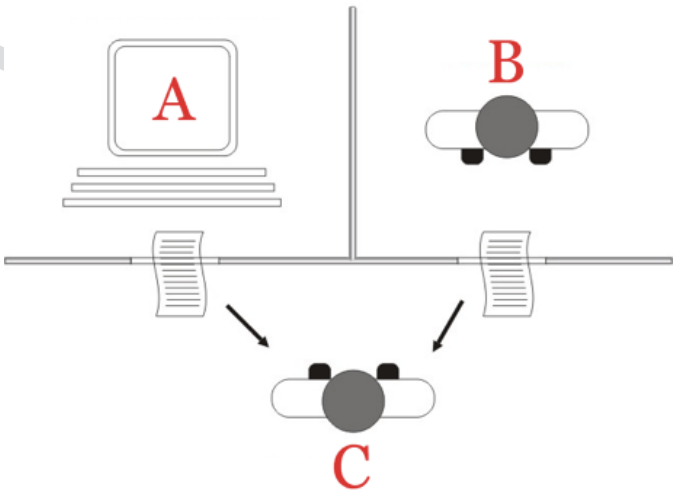


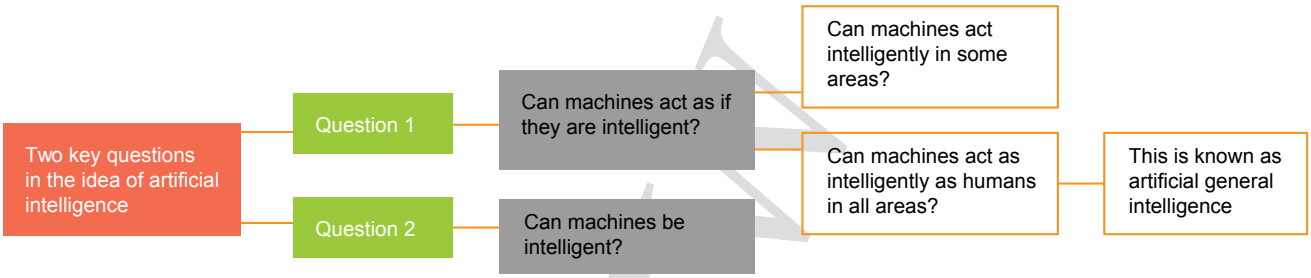
***Image 7**. Turing test diagram*
(credit: Juan Alberto Sánchez Margallo, Wikimedia Commons)

How can we know if something is intelligent? (#Methods and tools)

Two aspects of acting intelligently

Again, it is important to carefully separate two aspects of this question:

1) Can machines act intelligently in some areas? For example, in playing chess, in predicting weather, in piloting an airplane.
2) Can machines act intelligently in *all* areas? If they can, it means that machines can be as intelligent as humans not only in some things that humans do, but in all of the things humans do, in every walk of life. This question is also sometimes formulated like this: "*Can machines display general intelligence?*"



Area-specific intelligence

Chatbots provide a direct opportunity for Turing tests. Create a chatbot and have people converse with it. If they are unable to tell that they are having a conversation with a machine, then your chatbot passed the test. One of the first attempts to create a chatbot for this purpose was ELIZA, a "computer psychotherapist" designed by the MIT AI lab back in 1966. You can try having a conversation with ELIZA (http://psych.fullerton.edu/mbirnbaum/psych101/Eliza.htm) – just imagine you are visiting a psychologist and tell her about a problem you are experiencing.

Since this first awkward attempt, chatbots improved very quickly. In 1991, an annual competition was established for chatbots to try and pass the Turing test against a panel of human judges, called the Loebner Prize (https://aisb.org.uk). Eugene Goostman was the first to pass this test in June 2014. Eugene Goostman is a chatbot that simulates a 13-year-old Ukrainian boy. The conditions were quite strict: judges could ask any questions in a free unrestricted conversation, the conversation lasted for five minutes, and the chatbot was said to have passed the test if at least 30% of judges were convinced that they were talking to a human. Eugene managed to fool 33% of the judges (Veselov, 2014).

Can machines develop to the point where they will be able to solve all tasks humans can solve? (#Scope)

And of course you know about "personal assistants" such as Siri and Alexa. They are becoming more and more human-like. Google Duplex is a tool for making telephone calls to make appointments on your behalf. You tell your device that you want to book a table at a restaurant, for example, and it makes a call to confirm availability, check working hours and make a reservation. The program carries out a very realistic voice conversation with the person at the other end (Gewirtz, 2018).

***Image 8**. Chatbots are becoming more and more human-like*

So, we must admit that machines can act intelligently in specific areas. They are getting better and better. But let's go over to the second aspect of the question: can machines reach a point where they are human-like in *everything* we do? In other words, can machines display general intelligence?

KEY IDEA: There is no doubt that machines can act intelligently in some areas. Whether or not they can act as intelligently in all areas (in other words, display general intelligence) is not as obvious.

Artificial general intelligence

Here are some arguments in favor of the view that machines can indeed display general intelligence:

1) If we believe that the mind is a product of the brain (which is a belief shared by many) and that the brain obeys the laws of physics and chemistry, then there is no reason why we cannot recreate it. Assuming that technology will continue evolving, there is no obstacle to that.
2) If we view human reasoning as symbol manipulation that follows certain rules, there should be no doubt that we can teach computers to use these symbols and apply these rules. If symbol manipulation is all there is to the human mind, then we must be able to simulate it one day. If symbol manipulation is *not* all there is to the human mind, then what else is there?

KEY IDEA: If we accept that the mind is a product of the brain, it also seems that we must accept that artificial general intelligence will be possible in the future

Is human knowledge unique to humans? (#Perspectives)

Hubert Dreyfus (1929 - 2017) was one of the philosophers who rejected the idea that a machine can display general intelligence the way humans do. He claimed that the human mind is larger than just explicit manipulation of symbols following some set rules. To support this claim, he introduced the distinction between explicit thinking and implicit thinking. When we are solving a mathematical problem, for example, we use explicit thinking. We can formalize this process and teach it to others (and to computers). However, when we hunt a wild boar, for example, we use implicit thinking. It is unconscious and difficult to formalize. Hubert Dreyfus's argument was that most of human reasoning is implicit reasoning (intuition).

a. Alan Turing (aged 16) *(credit: PhotoColor, Wikimedia Commons)*

b. Hubert Dreyfus (a little over 16) *(credit: Jörg Noller, Wikimedia Commons)*

Image 9*. Alan Turing and Herbert Dreyfus*

There are skeptics like Dreyfus, but many thinkers today accept that there seems to be no reason to believe that machines will not be able to display general intelligence one day. Remember, they do not need to have a sense of humor or feel love, they just need to simulate feelings of love and display behavior that would suggest that they have a sense of humor. We may doubt machines will ever have minds, but surely they will be able to act as if they do?

Critical thinking extension

Is it morally permissible for us to build machines that will be superior to humans? (#Ethics)

The Turing test that we discussed in this lesson was designed as a test for artificial intelligence. The idea is that a computer's behavior is intelligent if it is indistinguishable from human behavior.

However, many AI researchers object to that. It is not the point of AI, they say, to imitate a human. When we build an airplane, for example, we are not trying to make it as similar as possible to a pigeon. And we are not judging its effectiveness by its ability to fool other pigeons into thinking it is one of them.

Would you agree that the point of AI is to build machines that will be capable of solving real-world problems better than humans? Would you agree that a machine does not have to think like a human to be intelligent?

If that is the case, what could we suggest as an alternative test for artificial intelligence?

If you are interested…

Alan Turing's contribution to our civilization is difficult to overestimate. He had a fascinating life full of triumph and tragedy. There is a popular movie based on his life, The *Imitation Game* (2014). Before watching it, I recommend reading the Wikipedia entry about this movie.

For a visual explanation, watch Alex Gendler's TED-ed video "The Turing test: Can a computer pass for a human?" (2016).

Watch the video "How the "most human human" passed the Turing test" (2018) on *Quartz*. It tells the flipside of the Turing test: the story of author Brian Christian, to be named the "most human human", who competed against artificial intelligence trying to prove to a panel of judges that he is indeed a human being.

Take-away messages

Lesson 2. There are two related questions about artificial intelligence: (a) Can machines act like they are intelligent and (b) Can machines be intelligent? It is important to not confuse the two. In this lesson, we focused on the first question. The most famous method used to answer this question is the Turing test. In this test, if a human having a conversation with another human being and a machine cannot tell the difference between them, then the machine is said to behave intelligently. There have been multiple attempts to build computers that would pass the Turing test and act intelligently in a certain area, for example, ELIZA, Eugene Goostman, Google Duplex. Many attempts have been successful. But this raises a further question: is it possible for machines to display general intelligence, that is, behave as intelligently as humans in all areas of human expertise? According to some, the answer is positive because there are no visible obstacles in recreating the structure of the human brain and the rules of human reasoning. Others (like Hubert Dreyfus) claim that most human thinking is implicit and difficult to formalize, so computers will not be able to imitate it. Yet others claim that imitating human thinking should not even be the goal of creating artificial intelligence.

Lesson 3 - AI: Artificial consciousness

Learning outcomes

a) [Knowledge and comprehension] What is the difference between artificial intelligence and artificial consciousness?
b) [Understanding and application] What are the arguments for and against the idea of artificial consciousness?
c) [Thinking in the abstract] How do we know that a machine *does not* have a consciousness?

Key concepts

Artificial consciousness, subjective experiences

Other concepts used

Chinese room (thought experiment), brain replacement scenario (thought experiment)

Themes and areas of knowledge

Theme: Knowledge and technology
AOK: Natural Sciences, Human Sciences

Recap and plan

We are investigating how technology can affect our knowledge of ourselves. In the previous lesson, we started looking at artificial intelligence and we agreed that there are two questions here that must be kept separate to avoid confusion: (a) Can machines act like they are intelligent? and (b) Can machines be intelligent? So far, we have been looking at the first question, and the answer is: yes, they can, at least in some spheres.

This brings us to the second question: can machines be intelligent? It's a much more difficult question where the Turing test will not be enough.

Can machines be intelligent?

Can technology know? (#Scope)

As you remember, there's a huge gap between machines *acting* intelligently and machines *being* intelligent. If you have a modern smartphone, you know that there's a whole range of things it can do: you can ask it (literally, using your voice) about the nearest restaurants with vegetarian food, and it will understand you, conduct a search and suggest some options. I don't see why such software can't be programmed to get offended when you say something insulting, to act like it is surprised when you say something out of the ordinary, and so on. You have a machine in your pocket that can act pretty intelligently. That's ok, you still know this is just a piece of metal and plastic, a well-designed *thing*.

However, what if I tell you that your phone *is* intelligent? That it can think and feel, be offended and surprised, perhaps even experience pain when you drop it on the floor? That it has a mind? This is where things get a little frightening, don't they?

Well, don't panic. First, let's agree on what "being intelligent" means.

What does "intelligent" mean?

According to Alan Turing, there is no difference between acting intelligently and being intelligent (Turing, 1950). This may seem a little weird at first sight, but the reasoning behind this claim is quite convincing:

1) We cannot observe someone's intelligence directly. We infer their intelligence from how intelligently they behave. This doesn't only apply to computers – we do that with each other.
2) Apart from inferring intelligence from behavior, there is no other way for us to tell if an entity is intelligent.
3) If an entity demonstrates intelligent behavior, it may or may not be intelligent, but our best option is to assume that it is.

***Image 10**. What is intelligence?*

KEY IDEA: According to Alan Turing, "being intelligent" = "acting intelligently"

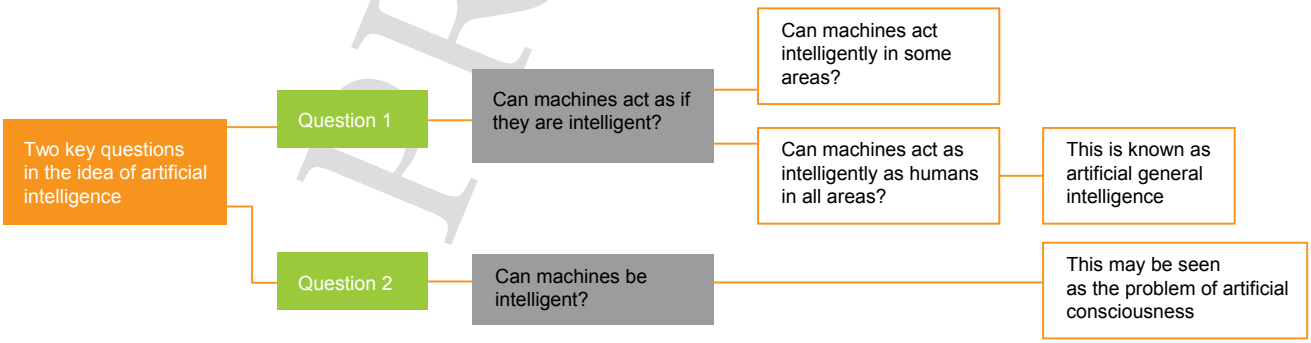
Not everyone felt comfortable with this reasoning. It feels weird to claim that my smartphone "has intelligence". This is because, subjectively, we experience this "something" within ourselves that produces intelligent behavior – our minds. The behavior of my smartphone may be very much like mine, but I have a mind and my smartphone doesn't. Right?

When I make a decision, I *experience* considering options and weighing possibilities, I *feel the pain* of disappointment if the outcomes are not what I expected. Although our behavior may be the same, computers don't *feel* or *experience* like I do. It is these subjective experiences that is emphasized by those who disagree with Turing's claim. By "being intelligent" they mean "having a mind", "having subjective experiences", "having mental states", "having consciousness".

So, can computers be intelligent in *that* sense? This now becomes a question of artificial consciousness. Artificial consciousness is the ability of computers to have subjectively experienced mental states. Let's agree that artificial intelligence means the ability of computers to *act* intelligently, but artificial consciousness means their ability to actually *be* intelligent.

Where is the line between a mind and a thing? (#Perspectives)

KEY IDEA: Those who disagree with Alan Turing suggest that to be intelligent, one needs to have subjectively experienced mental states (consciousness). The question then becomes, can machines have consciousness?



John Searle's "Chinese room"

Arguing against the idea that computers can have minds, in 1980 John Searle came up with a thought experiment that he called the "Chinese room" (Searle, 1980). It has been widely discussed ever since.

Suppose that AI scientists have succeeded in designing a computer that acts as if it understands Chinese. The software takes Chinese characters as inputs, processes them and produces sequences of Chinese characters as outputs. Suppose also that this computer successfully

passes the Turing test: Chinese-speaking humans interacting with it are convinced that they are conversing with another human Chinese speaker.

Now, imagine Searle himself sits in a closed room where he has a book with an English version of the computer program, papers and file cabinets to record and store information, and pencils and erasers to write down his answer. He receives Chinese characters through a slot in the door, processes these characters according to the instructions in the book, and writes his output on a card that he pushes back through the slot in the door. Essentially, this is doing what the computer does, only manually.

How can thought experiments be helpful in gaining knowledge? (#Methods and tools)

Searle claims that in this thought experiment there is no essential difference between himself and the computer that follows instructions step by step and spits out an output that is interpreted by human beings as intelligent behavior. But just like Searle doesn't understand a word of Chinese, the computer would not understand Chinese either. And there is nothing in that room that can be said to understand Chinese. Since the computer does not understand Chinese, it does not have a mind and it is not intelligent.

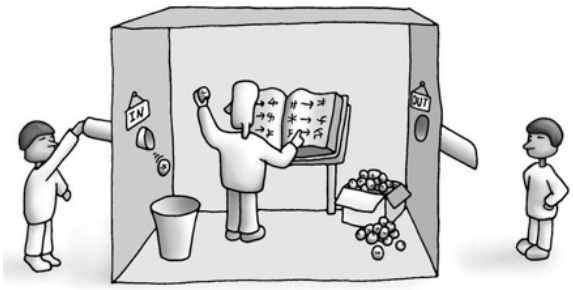


***Image 11**. Chinese room thought experiment*

Arguments against "Chinese room"

AI scholars have made multiple attempts to refute the argument, generating some interesting debates.

One reply was that the mind in the Chinese room is not the man, but the whole system: the man plus the papers and file cabinets and pencils and erasers. The man does not speak Chinese, but the room does.

How do we know that we have a mind? How do we know that someone else has a mind? (#Methods and tools)

Another reply is *the brain replacement scenario*. Searle says that a computer program (or a machine) cannot be conscious no matter how closely it simulates the human brain. Imagine that scientists have invented a tiny computer that simulates the function of an individual neuron. They start gradually, one by one, replacing the real neurons in your brain with these simulated devices. If they replace one neuron, that would probably do nothing to your consciousness. But what happens when scientists continue replacing more and more neurons in your brain? According to Searle, a completely artificial brain must not have consciousness, therefore you must lose conscious control at some point during this process. Imagine that part of your brain has been replaced with these artificial neurons. Your teacher asks you "Do you believe that machines can have minds?", and you want to shout "No, never!", but much to your dismay you hear your own voice saying "Yes, definitely". Critics find this scenario weird; they say that there will be no such point where conscious awareness is replaced by automatic, mindless reactions. Therefore, conscious awareness will remain a property of the fully artificial brain.

***Image 12**. In a hypothetical scenario parts of the brain are replaced by artificial neurons*

Conclusion

We don't have any satisfactory answers yet. The idea of a conscious machine is somehow counter-intuitive. It goes against our subjective experiences that a "thing" can have a mind just like our own. At the same time, we know that the brain theoretically can be reproduced. If things do not have minds, then an artificial brain will not have a mind, either. But then it is unclear what "a mind" is. If it is not entirely a product of the brain, then what is it and where does it come from? Unless we answer this question convincingly, we will need to accept that things can have minds.

KEY IDEA: The idea of a conscious machine is counter-intuitive because "a thing cannot have a mind". But then it is very difficult to explain what else is there in the mind that cannot be reduced to the thing.

Critical thinking extension

How do we know that a machine *does not* have a consciousness?

Imagine that the day has come when we have built an android that is indistinguishable from a human. The android acts like a human being in everything it does. For example, when it touches a hot surface, it pulls back its hand and screams as if it was in pain. Now, my question about this android is: is it human? Should it be given the same rights as human beings?

It probably depends on whether or not the android has consciousness. Does it experience pain or does it merely act as if it is experiencing pain?

Imagine this android is *you*, and you do experience pain and have consciousness, but people around you are convinced by John Searle's arguments and believe that you are merely a thing. How do you prove them wrong?

If you are interested…

Watch Joscha Bach's TED talk "From Artificial Intelligence to Artificial Consciousness" (2016) – insightful, though slightly on the technical side.

Watch the video "These self-aware robots are redefining consciousness" (2019) on the YouTube channel *Seeker*. This video is about a research lab that tries to build self-aware robots and their latest achievements.

Watch David Chalmers's talk "Artificial consciousness" (2016) on the YouTube channel *Serious Science*.

Watch the video "The Chinese room experiment – The hunt for AI" (2015) on the YouTube channel *BBC Studios*.

If you have not had these lessons already, you might want to have a look at lessons about "qualia" in the chapter "Knowledge and understanding". These lessons have many concepts and thought experiments that are related to our discussion of artificial consciousness.

Take-away messages

Lesson 3. We have seen that machines can act as if they are intelligent. Some even think that machines can display general intelligence, that is, they can *seem to be* as intelligent as humans in every walk of life. But the next question is, can machines be intelligent? Many thinkers assert that acting as intelligently as a human does not mean being intelligent. Many thinkers assert that it does. What is usually meant by intelligence in this context is "subjective experiences", "mental states" or "consciousness". So, this debate can be more accurately described as a debate over artificial consciousness. John Searle with his thought experiment "Chinese room" proposed that a machine cannot be intelligent even if it is an exact copy of the human brain. However, some counter-arguments were proposed too, for example, it is not clear where intelligence (consciousness) disappears when a human brain (in a hypothetical scenario) gradually turns into an artificial brain.

Lesson 14 - Redefinition of art

Learning outcomes

a) [Knowledge and comprehension] How did art redefine itself historically in the process of development?
b) [Understanding and application] In what ways does digital technology trigger a redefinition of art?
c) [Thinking in the abstract] How will technology change our understanding of authorship and originality in art?

Key concepts

Redefinition of art, authorship, originality

Other concepts used

Realism, impressionism, modern art, photography, photocopying, irreproducibility, AI-generated texts

Themes and areas of knowledge

Theme: Knowledge and technology
AOK: The Arts

Recap and plan

We are investigating the role of technology in obtaining knowledge about the world. We have looked at phenomena such as computer simulations and Big Data. We have decided that these phenomena might have the potential to trigger revolutionary changes in natural sciences, human sciences and history. We have also looked at the relationship between knowledge and technology in mathematics, and our conclusion was a little different. While experimental mathematics changes the landscape of day-to-day work of a mathematician, the fundamentals of the traditional deductive method of reasoning have not been challenged.

There's one area of knowledge left to consider – the Arts.

In this lesson, I will briefly look at the phenomenon of redefinition of art. I will claim that art develops by redefining itself in response to some major challenges or "uncomfortable questions". Some of these challenges come from newly emerging technology, as can be seen from examples such as the invention of photography, then photocopying, and now – digital devices.

How can art be defined (if it can be defined at all)? (#Scope)

Art develops by redefining itself

Art redefined itself every time it was challenged by something and went through a period of crisis. To illustrate, I will give two examples related to technology.

***Image 40**. Andy Warhol's "Campbell's Soup Cans"*

At some point in the 19th century, *realism* in art was the dominating trend. The purpose of art was to represent reality as it is, as accurately as possible. Artistic skills were of great value. It took years of traditional education in an art academy to develop these skills. Then along came photography. It became possible to click a picture of a landscape and get a realistic representation of it – without years of training, without weeks of work. Art was in crisis. To survive, it had to redefine itself, and it did. The impressionist movement started deviating from academic standards, emphasizing the importance of capturing the artist's impression of reality rather than reality itself. Photography could not compete with art anymore because a photograph cannot capture your subjective impression. Art saved itself through redefinition.

Another example was the invention of photocopying. For a long time, uniqueness had been considered one of the defining characteristics of a work of art. To be unique meant to be irreproducible. The original copy of the *Mona Lisa* is displayed in the Louvre Museum, and to see it you need to physically go there and stand in line for a couple of hours. With the invention of photocopying and mass print production, art faced some uncomfortable questions. Does a good photocopy of the *Mona Lisa* have the same artistic value as the original? How can art keep its uniqueness and protect itself from being massively reproduced? Then along came Andy Warhol and, with his famous *Campbell's Soup Cans* (1961-1962), redefined art. The *Soup Cans* were as reproducible as one could possibly imagine. With this work, Warhol claimed that art can and should be reproducible. Modern art embraced this idea. Warhol's *Soup Cans* is a great work of art precisely because it redefined art, eliminating irreproducibility from its definition.

These are just a few examples. But clearly the emergence of digital technology could not go unnoticed, and art in its development must have reacted to it.

How does art react to challenges presented by new developments in technology? (#Methods and tools)

How digital technology challenges art

Just like photography challenged realism in art and photocopying challenged the idea of irreproducibility, digital technology raises several questions that challenge the very essence of art. For example:

1) If a work of art is produced by a machine, does it still count as art?
2) Who should be credited for a work of art? Suppose an artist has created an algorithm that draws on a canvas, then the algorithm created an image using a graphic software. Does credit go to the artist? To the algorithm? To developers of the graphic software?
3) What is the nature of originality? Can computer-produced art be called "original"? If not, what exactly makes human-produced art more original than computer-produced art?

***Image 41**. Can robots create art?*

It seems we live in exciting times because, just like after the invention of photography years ago, art will have to embrace these new developments like it usually does – by redefining itself. But the questions are tough, so I wonder if art will survive this time.

Can art be produced by a computer? (#Perspectives)

KEY IDEA: When art is challenged by a technological innovation, it redefines itself

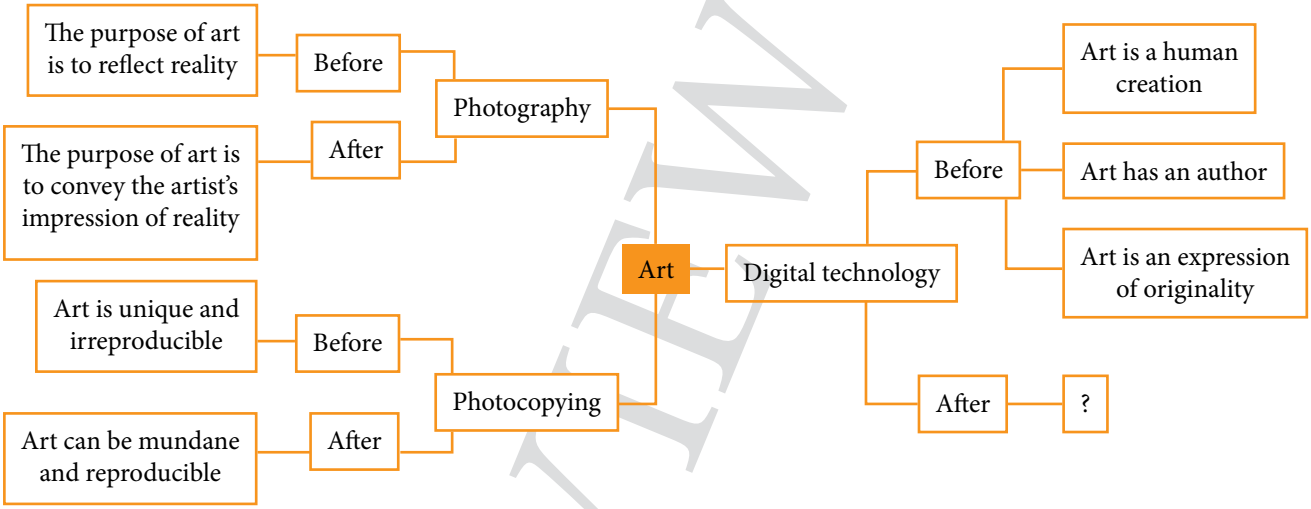
Example: Harry Potter and the Portrait of What Looked Like a Large Pile of Ash

Let me give you one example to illustrate all the controversy that modern technology can create in terms of defining or redefining art.

Have you read the AI-generated chapter of *Harry Potter*? If not, please do me a favor and read it (see "If you are interested" below)! It comes from the tech company *Botnik Studios* and the name of the chapter is "Harry Potter and the Portrait of What Looked Like a Large Pile of Ash".

An algorithm was trained on all seven *Harry Potter* books by J.K. Rowling. In doing so, the algorithm picked up the most commonly used words and word combinations, characteristic word order, the use of suffixes, and so on. After being trained, the algorithm produced a text of its own – what Rowling could write. This did not look too meaningful (algorithms are not that good yet!), but then a team of human authors took the product and cleaned up each sentence a little. You can see the result of this work for yourself.

At times it is funny, at times it is surprisingly creative and at times it is gibberish. But you will probably agree that it is a good read and time well spent. And it does have the style and the vibe of the original novels.



Two uncomfortable questions

Using this example, let me come back to the uncomfortable questions that art faces now that digital technology is being developed. I will try to formulate two questions that I believe are most crucial.

Question 1: Can this AI-generated chapter of *Harry Potter* be considered a work of art? It ticks a lot of boxes. It seems like it was intended as such. It can certainly be perceived as such, at least by some audiences. And in itself, I would claim that this is much better quality work than some of the human-generated pieces of literature I have seen.

Question 2: Who is the author of this chapter? There is little doubt that the algorithm is a creation of the coders. But once created, the algorithm works all by itself. The algorithm was trained on *Harry Potter* books by J.K. Rowling. The whole purpose was to write as much like J.K. Rowling as possible. So, can we claim that she is the author of this chapter, or at least one of the authors? She might not even know that this chapter exists.

What does it mean to be "original" in art? (#Perspectives)

It looks like originality and authorship will be the dimensions that art will have to redefine in itself, due to the newly emerged technology.

***Image 42**. Uncomfortable questions*

Critical thinking extension

Two more uncomfortable questions

Two more uncomfortable questions about the AI-generated *Harry Potter* chapter – these two are somewhat more general than the ones we have already discussed in the lesson, although they are also related to originality and creativity.

Question 3 (follow-up on authorship): Can art be produced by a computer? If you believe that the author of the chapter, at least partly, was the computer algorithm, then you must also accept that computers can create art. But if that is so, then what is the role of humans? Is this the end of human art?

Question 4 (follow-up on originality): What is the nature of originality in art? A common answer to the previous question is "No, art cannot be produced by a computer because computers can only follow an algorithm while humans can produce original creations". Then the question is, how exactly is a human original creation different from a computer implementing an algorithm?

I am very curious: what are your answers to the uncomfortable questions raised in this lesson?

If you are interested…

Check out the work of *Botnik*, a "machine entertainment company" as they call themselves (https://botnik.org/). Their *Harry Potter* chapter is available on their website.

Botnik

Read Janelle Shane's blog post "The neural network generated pickup lines that are actually kind of adorable" on the website *AI weirdness*. She does a lot of funny stuff with AI. Teaching AI to generate pick-up lines is just one of her projects.

Have a look at the book *Introducing Postmodernism: A Graphic Guide* by R. Appignanesi and C. Garratt (2003). This book gives an excellent overview of the history of how art redefined itself in response to various challenges. Fun to read and very insightful, it is highly recommended if you want to understand art better.

AI Weirdness

Take-away messages

Lesson 14. In this lesson, we started looking at the role of technology in art. The history of development of the Arts as an area of knowledge is a history of art redefining itself in response to periods of crisis or challenges raised by technological and other developments. We have considered two examples of this. First, the invention of photography which resulted in art redefining itself from capturing reality to capturing an impression of reality. Second, the invention of photocopying and mass production which resulted in art rejecting the idea that irreproducibility should be one of its defining characteristics. Surely, modern digital technology also presents a challenge that art needs to respond to by redefining itself. We considered the example of an AI-generated chapter of *Harry Potter* and discussed some uncomfortable questions that art needs to answer today: (1) is it possible for a machine to produce art? (2) what is the role of human skill in art? (3) who should be credited for a work of art produced with major assistance from a computer? and (4) what is the nature of human originality?

UNIT 3 - Bias in personal knowledge

Exhibition: a turbulence map	156
Story: Senate Bill 464	157
Lesson 1 - Bias	158
Lesson 2 - Personal experience	161
Lesson 3 - Darwinian evolution of personal knowledge	165
Lesson 4 - Analogy analysis	169
Lesson 5 - Cultural experience	173
Lesson 6 - Memes and Universal Darwinism	177
Lesson 7 - Heuristics	181
Lesson 8 - Implicit bias and bias self-awareness	185
Lesson 9 - Bias reduction	189
Lesson 10 - Compos mentis	193
Back to the exhibition	197

UNIT 3 - Bias in personal knowledge

You may remember from Unit 1 ("Knowledge of knowledge") that there is a distinction between personal knowledge and shared knowledge. These terms are quite transparent: personal knowledge is something belonging to you as an individual, while shared knowledge is something common to sizeable groups. Shared knowledge and personal knowledge are overlapping circles on a Venn diagram. Some of your personal knowledge coincides with that shared by other people, but another part of your personal knowledge is unique to you.

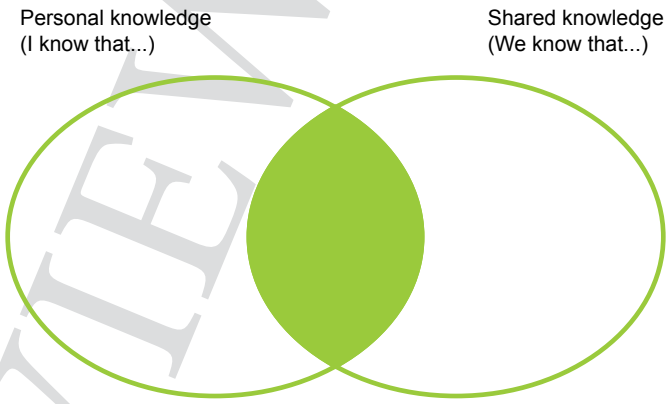


***Image 1**. Personal knowledge and shared knowledge: how they are related*

The good thing about bias is that, although every individual is biased, collectively we can keep these biases in check and overcome them. In a series of independent replications, conclusions of one scientist may be validated by other scientists. In a jury court, opinions of the jurors may be compared and discussed. Scientists may have different explanations for an observed phenomenon, but through testing and replication some explanations are eliminated and some retained. In other words, biases are abundant in the realm of personal knowledge, but not so much in shared knowledge. As a rule, shared knowledge is much less biased than personal knowledge.

Can biased personal opinions be valuable for developing shared knowledge? (#Perspectives)

The bad news is that shared knowledge can also be biased. Biased shared knowledge is probably more disastrous than biased personal knowledge simply because we trust it more. Additionally, it is much more difficult to identify the bias and eliminate it when it is the whole of humanity that is biased. In other words, although biases in shared knowledge are less numerous, they are more impactful.

	How many biases are there?	How impactful are they?
Personal knowledge	A lot!	They affect only you
Shared knowledge	Not so many	They affect everyone!

KEY IDEA: Biases in shared knowledge are less numerous, but they are more impactful

In this unit, we will consider biases in personal knowledge. On the surface, the problem may seem simple: just check your personal knowledge against shared knowledge and get rid of your bias! However, we cannot just dismiss personal knowledge as something inferior to shared knowledge. After all, as a knower, your personal knowledge is all you have access to. A belief that you retrieve from your personal knowledge can either come from the area that overlaps with shared knowledge or from the area that is uniquely yours. How do you know which area it comes from?

Can we know if our personal knowledge is biased without checking it against shared knowledge? (#Methods and tools)

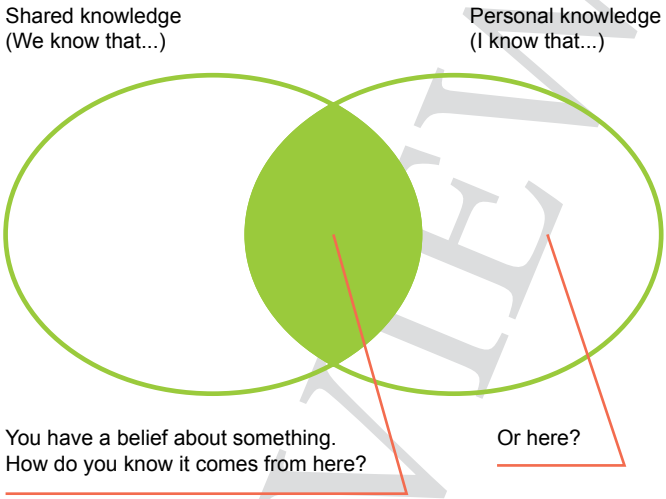


***Image 2**. Where does your belief come from?*

The knowledge that you are directly in touch with and that you use on a daily basis is your personal knowledge. For this reason, personal knowledge is worth considering on its own before we move on to biases in shared knowledge.

Exhibition: a turbulence map

In front of me is an aviation weather forecast chart (for simplicity I will call it a turbulence map).

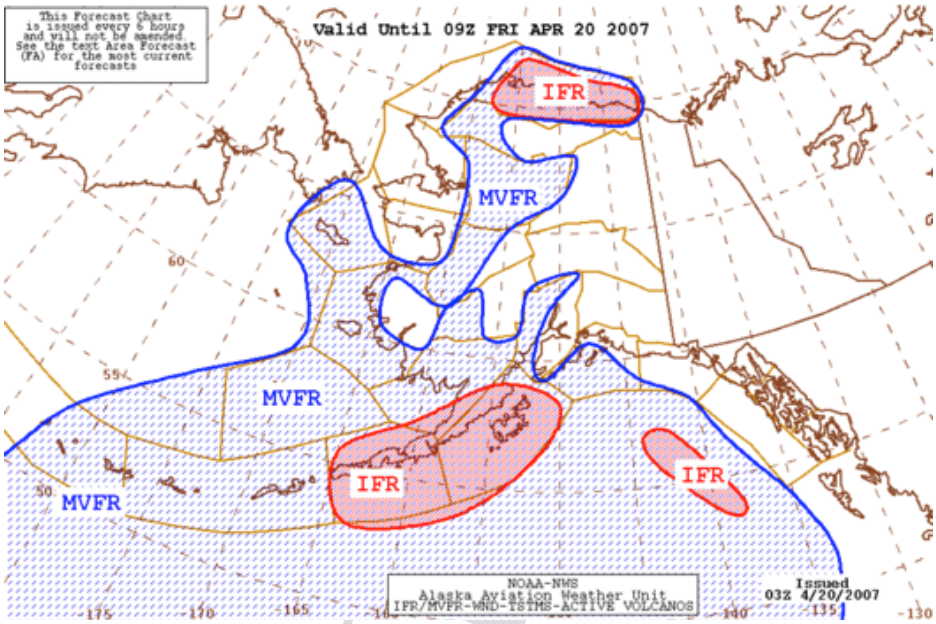


***Image 3**. Aviation weather forecast chart (turbulence map) (credit: Wikimedia Commons)*

Such maps show you the areas where turbulence is more likely to occur when you are travelling by air. These maps (among other sources of information) are used by pilots to try to make your flight smoother when they are navigating.

I am a nervous flyer. I have a complicated relationship with turbulence. It is pretty unfortunate for someone who works in an international setting and needs to travel a lot.

At some point when it became really irritating, I started educating myself. I read articles and watched videos that explained turbulence and analyzed past airplane crashes. I discovered that a lot of my beliefs had been inaccurate and misleading. First of all, I used to think that turbulence can cause airplanes to crash. Now I know that airplanes are designed so that they can withstand turbulence more than two times stronger than anything commercial flights are likely to encounter. I used to think turbulence was the most dangerous part of the flight. Now I know that you are more likely to be harmed while you are on the tarmac than when you are experiencing turbulence mid-air. I used to think air travel was a risky option. Now I know that statistically I am much more likely to die in a car on the way to the airport.

Has it helped? No. Every time turbulence kicks in, I still grab the armrest until my knuckles turn white. In reality, I should be doing that in taxis, not in planes! My conscious brain knows that, but my body seems to refuse to listen.

I still check "turbulence maps" before flying. The abundance and accessibility of such maps online gives me a hint that I am not alone. It appears as though there are many more nervous flyers out there who misinterpret the danger of planes (relative to other means of travel), whose logical brain cannot override the rest of their brain, whose expectations, perceptions and attitudes to air travel are all biased because of this complicated relationship with turbulence.

The truth is, if your seatbelt is fastened, turbulence is not dangerous. My beliefs and perceptions, however, systematically deviate from this truth in the direction of misinterpreting various aspects of air travel as more dangerous than they really are.

Story: Senate Bill 464

The year 2019 in the USA saw an unusual precedent in legislation: Senate Bill 464 made it mandatory for doctors and nurses in California to undergo eight hours of implicit bias training and testing periodically (every 2 years).

This is probably one of the first times when the concept of implicit (unconscious) biases made its way into legislature.

***Image 4**. There are racial differences in the chance of death from complications at childbirth*

This bill was "inspired" by some disturbing research findings that showed that, although there was a decrease in the overall number of women who died giving birth in California, black women were still 3 or 4 times more likely to die from complications at childbirth compared to white women. Additional research into this issue showed that roughly half of surveyed medical professionals believed myths and shared misconceptions about racial differences in tolerating pain. For example, they believed that black patients can "endure more pain" and have "thicker skin". Such biases created a situation where, when an expectant black mother claimed she was in pain, doctors underestimated the severity of her condition and did not respond appropriately. Obviously, the medical professionals were entirely oblivious of this bias that they had. This research was conducted in 2016. While it is quite hard to believe that such racial biases are so widespread in the 21st century, we cannot simply attribute this to "bad doctors". These biases are implicit – they occur without the conscious awareness.

The bill requires medical professionals to go through training that teaches them to identify their own implicit biases and consciously counteract them. This is an attempt to reduce discrimination by targeting our own unconscious minds.

You can read more about the Bill in the article "These California bills would train nurses, judges and police how to spot their own biases" in *Los Angeles Times*.

The full text of the bill can also be found online, its name is SB-464, California Dignity in Pregnancy and Childbirth Act.

Lesson 1 - Bias

Learning outcomes

a) [Knowledge and comprehension] What is bias?
b) [Understanding and application] What are the key examples and non-examples of bias?
c) [Thinking in the abstract] How can bias be separated from similar knowledge concepts (such as prejudice, misconception or superstition)?

Key concepts

Bias, systematic deviation, opinion, perspective, mistake

Other concepts used

Stereotype, prejudice, misconception, superstition, decision-making

Themes and areas of knowledge

Theme: Knowledge and the knower

Plan

In this lesson we will define bias and think about examples and non-examples of bias. In line with the purpose of this unit, the focus will be on bias in personal knowledge. Just to remind you, bias in personal knowledge may be assessed against shared knowledge. If we want to know if our personal belief is biased or not, we can compare it to the accepted, well-established beliefs on the same subject matter that we have collectively agreed upon.

Shared knowledge, of course, can also be biased, but that will be the focus of the next unit.

What is bias?

Is it true that we are much more biased than we could possibly imagine? (#Scope)

As much as I would like to think of myself as an open-minded, unprejudiced, impartial and just individual, I know that I am not one (are you?). Growing up, I was influenced by a variety of factors and exposed to a variety of experiences. In all probability, these experiences have caused me to have certain biased beliefs. Worst of all, I am probably biased in ways that I am not even aware of.

I will define bias as a *systematic deviation from the truth*.

When I say "*deviation*", I imply that there exists a correct answer (belief, decision) and that the answer (belief, decision) we are dealing with does not match this correct one. This is important because we can identify a bias only if we know the correct answer. If we do not know what the correct answer is, or if we cannot at least assume the correct answer beyond a reasonable doubt, there is no point in talking about bias.

When I say "*systematic*", I mean a deviation that is not random. In other words, it is leaning consistently towards one direction rather than various directions at various times. For example, suppose you are measuring the width of your bed with a measuring tape. You carry out the measurement 10 times. Every time you will get slightly different readings, both higher and lower than the real width of your bed. This is an example of *measurement error*, but this is not a *bias*. A bias occurs when, for some reason, the measurement deviates systematically in one direction. For example, suppose the measuring tape itself is flawed – you washed it accidentally in the washing

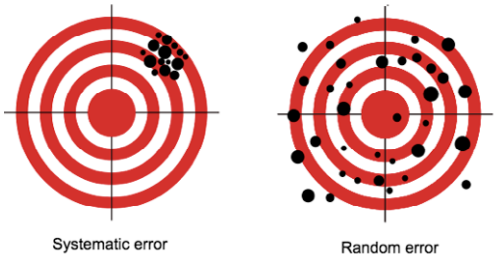


***Image 5**. The difference between systematic error and random error (credit: Wikimedia Commons)*

machine and it shrank a little, resulting in each inch section being a little shorter than it is supposed to be (I am now assuming that it is a cloth measuring tape, not a metal one… why would you put a metal measuring tape in a washing machine?). In this case, no matter how many times you carry out the measurement, you will always underestimate the width of your bed. This is bias.

KEY IDEA: Bias is a systematic deviation from the truth

Sources of bias

Since the deviation is systematic, it is usually the case that the deviation is *caused* by something, in other words, that there is a source of bias. In my turbulence example, overestimating the dangers of air travel is caused by my fear of turbulence. It also probably means that whenever there is bias, we can identify one or several factors that make it happen.

Theoretically:

- If we can eliminate the source, the bias will disappear
- If we know the source, we can predict the bias (for example, knowing that a person has a fear of turbulence means that we can probably predict that they will overestimate the dangers of air travel)

Is there any way to know what causes our personal bias? (#Methods and tools)

There are many possible sources of personal bias. Some of them are linked to our identity (cultural, political, gender). Some are linked to our personal experiences (having survived through certain difficulties, having witnessed certain events). Arguably, every human being has a different background and that could determine how (in what way) they are biased.

The important take-away message here is that biases are systematic because they are systematically affected by a certain source and, at least theoretically, these sources can be identified and dealt with.

Bias versus other concepts

To understand a concept, it is always useful to separate it from (misleadingly) similar concepts by answering the question "What is it not?"

We have defined bias by stating what it is. Let us now try to delineate it from a variety of other concepts that it can be easily confused with.

Bias is not the same as opinion. Opinions are possible when there is no single truth. For example, it is my opinion that restaurant A is better than restaurant B. Airplanes falling because of turbulence cannot be my opinion because we do know that this is false. Since we have access to a pretty unambiguous truth in this case, opinions are no longer a thing – there are either beliefs that correspond to the truth or ones that don't.

Bias is not the same as perspective. Again, perspectives are possible when the truth is complex and when multiple interpretations of the truth are possible. For example, there may be various historical perspectives on events of the past. There can be various angles of looking at those events, and often there is no way to prefer one perspective over another. For this reason, perspectives are very valuable (the more the better!). By contrast, in my turbulence example, the truth is pretty straightforward. Another difference is that, when you are presenting a perspective, you are presenting it honestly as one of several possible angles in looking at a situation. You acknowledge the existence of other angles. When you are biased, you are trying to pass your bias off as the truth (and you actually believe it to be the truth).

Is it possible for biases to be accepted as valuable perspectives? (#Perspectives)

Bias is not the same as a mistake. It is a particular type of mistake – a systematic one. If I ask a child who has never travelled by air if turbulence can bring down airplanes, they may say yes. It would be a mistake but not a bias. If you ask someone like me (before they educated themselves with loads of articles and videos), they will say yes because they are afraid of turbulence. They will answer multiple other questions with similar mistakes – for example, they will overestimate the likelihood of turbulence occurring, the psychological effect it has on airline pilots, and the number of turbulence-related accidents in the past. All of their answers will be biased in the same direction, driven by one source - their underlying fear of turbulence.

Critical thinking extension

Now that we are clear with the definition of bias and with some of the things that bias is *not*, can we name some examples of phenomena that may be categorized as instances of bias in personal knowledge?

To what extent can we claim that personal bias penetrates every aspect of our lives? (#Scope)

Here are some of the phenomena that we are going to consider further on in this unit:

1) Biased perception (for example, susceptibility to certain perceptual illusions)
2) Stereotypes
3) Prejudice
4) Biased decision-making (for example, selecting risky options when it is not logically warranted)
5) Misconceptions (biased understanding of certain ideas, not just a mistake but a systematically incorrect understanding driven by a false belief)
6) Superstitions (stubborn beliefs in supernatural influences despite counter-evidence)

Do you think all of these phenomena fit our definition of bias equally well? Would you add any other phenomena to the list?

If you are interested…

When a meteorologist talks about bias, it is worth listening to (I would know, both of my parents have degrees in meteorology). J. Marshall Shepherd's TED talk "3 kinds of bias that shape your worldview" (2018) is a good place to start.

Take-away messages

Lesson 1. Bias is a systematic deviation from the truth. This definition implies two things: (1) there exists a certain standard that we may accept as the correct answer or the truth, (2) the deviation from this standard is not occasional and random, but systematic (consistent and always in the same direction). For this reason, opinions, perspectives and mistakes are all non-examples of bias. Since biases are systematic, it must be the case that they are (systematically) influenced by some factors. Such factors are known as sources of bias and they can originate from your personal experiences, your culture, your identity, and so on.

UNIT 4 - Bias in shared knowledge

Lesson 1 - Naïve theories 201

4.1 - Bias in Natural Sciences 205

Exhibition: Refracting telescope 205
Story: Discovery of Neptune 206

Lesson 2 - Demarcation problem 207
Lesson 3 - Falsifiability 212
Lesson 4 - Scientific progress 217
Lesson 5 - Underdetermination of scientific theories 222
Lesson 6 - Theory-laden facts 227
Lesson 7 - Verisimilitude 232
Lesson 8 - Paradigm shifts 236
Lesson 9 - Incommensurability 240

Back to the exhibition 244

4.2 - Bias in History 245

Exhibition: British History for Dummies 245
Story: The Battle of Waterloo 247

Lesson 10 - Historical interpretation 248
Lesson 11 - Historical perspectives 253
Lesson 12 - Historical objectivity and historical facts 257
Lesson 13 - Historical objectivity and rival interpretations 262
Lesson 14 - Historical objectivity and ethics 266
Lesson 15 - Heteroglossia (in theory) 270
Lesson 16 - Multiperspectivity (in practice) 274

Back to the exhibition 278

4.3 - Bias in Mathematics 279

Exhibition: A FIFA football 279
Story: George Dantzig's homework 280

Lesson 17 - Proof 281
Lesson 18 - Axiomatic systems 286
Lesson 19 - Discovered or invented? Truth in mathematics 290
Lesson 20 - Consistency 295
Lesson 21 - Mathematical realism 299

Back to the exhibition 303

Lesson 22 - Overview: bias in Mathematics, Natural Sciences and History 304

4.1 - Bias in Natural Sciences

Natural sciences study the objectively existing world of material things. Therefore, the main point of reference when identifying bias in natural sciences is a deviation from the reality of things. If our beliefs about things deviate from how things really are, then we are dealing with bias. This seems simple enough.

But it's not that simple.

Examples of bias would include all scientific misconceptions and faulty theories that used to be accepted at one time but were later replaced by better theories. The ether theory in physics, the phlogiston theory in chemistry, Lamarckian views of evolution of species, the geocentric model of the world – all these and many other ideas used to be widely accepted but have been replaced.

The key questions that can be asked in this respect are:

- Why do we accept incorrect theories in the first place? Can't we see that these theories have no correspondence to the reality of things?
- What is the best way to establish correspondence between beliefs and reality?
- How do ideas replace each other in sciences? As old ideas get replaced by newer ones, are we getting closer to "the truth"?

To answer these questions, we must consider several key concepts:

- Demarcation criteria
- Falsifiability
- Underdetermination of scientific theories
- Theory-laden facts
- Verisimilitude
- Paradigms and paradigm shifts
- Incommensurability

As all of these concepts are closely linked with bias, they will be the focus of our discussion in the next several lessons.

Exhibition: Refracting telescope

In front of me is a simple refracting telescope. A device that should enable me, as its name suggests, to see (*scopein*) far away (*tele*). Above me is a vast night sky. I want to use my device to see what's out there, to get to know the Universe I live in.

But I have a doubt: will my telescope show me the truth? Will it show me the Universe as it is, without distortions? Can I trust it to be my guide? Will what I see through my telescope be distant celestial objects that are floating out there, or will it be some properties of the telescope itself that I mistake for stars and planets? This can certainly happen if there are dust particles on the lens. I can clean the dust, but how do I know the telescope does not have any other inherent biases? What if the lenses filter something out? What if they give me a distorted image with incorrect angles? If I fail to see something that is actually out there, I can probably live with that. But what if I see something that is *not* out there? That would be very disappointing and misleading.

I usually trust something that I see with my own eyes – but can I have the same amount of trust in something I see through a strange device invented by a scientist?

***Image 8**. Refracting telescope (credit: Mike Peel, Wikimedia Commons)*

My refracting telescope is the simplest of them all. It uses lenses to form an image. The lens bends (refracts) the light from a distant object and focuses it. It can gather more light than the human eye can manage. A telescope working on the same principle was used by Galileo Galilei in his observations (back then, you couldn't just order one on Amazon, so Galilei had to actually construct his own).

Since then, there have been many modifications and all sorts of telescopes working on different principles: reflecting telescopes that use mirrors to collect and focus light; X-ray and infrared telescopes; radio telescopes that have antennas that collect radio waves and microwave radiation; gravitational wave detectors; space telescopes such as the Hubble Space Telescope that is orbiting the Earth.

If I don't trust something as simple as my two-lens telescope to provide an accurate picture of reality, how can I trust something as complicated as a gravitational wave detector?

How do I know that my telescope is not biased?

Story: Discovery of Neptune

When the telescope was invented, scientists meticulously observed the sky and discovered planets in the Solar System that were not visible to the naked eye. For example, the year 1781 saw the discovery of Uranus.

This was also the era of Newtonian mechanics. Newton's (and Kepler's) equations described the motion of celestial objects and explained it by the influence of gravitation. It looked very promising because the planet trajectories that astronomers observed coincided with those predicted from Newton's equations.

But not for Uranus. As astronomers were observing it since its discovery, its orbit did not match precisely with what was expected of it. This could mean that Newton's equations were wrong. Or it could mean that Uranus was influenced by another force that the astronomers were not accounting for. Could this force be gravitational pull from another, unknown planet? If so, then it could be possible to look at Uranus's deviations from the predicted trajectory and use the equations to calculate where this force should be coming from.

In 1845, astronomers Le Verrier and Adams independently carried out calculations to determine the position of this hypothetical unknown planet. In 1846, astronomers at the Berlin Observatory pointed their telescopes at the location predicted by these calculations and voila! They saw a planet that no one had noticed before - Neptune.

In other words, Neptune was mathematically predicted before it was directly observed through a telescope. The magic of this story is that a planet was discovered "with the tip of a pen", from the comfort of a scientist's desk.

To be fair, analysis of old documents reveals that Neptune had actually been observed many times before but had not been recognized as a planet. For example, Galileo observed it in 1612 but mistook it for a distant fixed star. Some great astronomers of the past didn't recognize Neptune even when they looked at it through a telescope, while Le Verrier and Adams did not even have to look at it to know that it's there. Weird, isn't it?

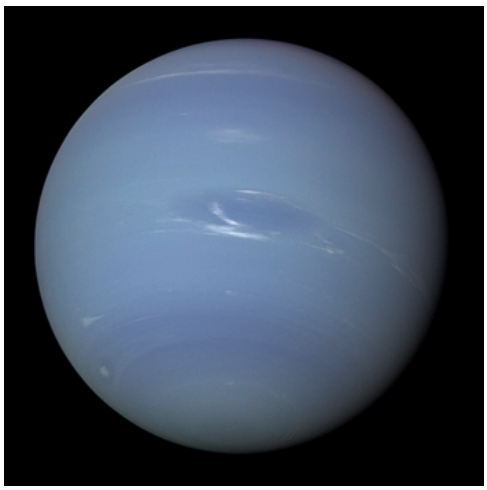


***Image 9**. A photograph of Neptune taken by the Voyager 2 spacecraft in 1989 (credit: Justin Cowart, Wikimedia Commons)*

Lesson 2 - Demarcation problem

Learning outcomes

a) [Knowledge and comprehension] What is a demarcation criterion?
b) [Understanding and application] Why is demarcation based on empirical verification of statements logically flawed?
c) [Thinking in the abstract] How can we draw a line between science and non-science to ensure that what is categorized as science guarantees knowledge that is beyond a reasonable doubt while what is categorized as non-science doesn't?

Key concepts

Demarcation problem, demarcation criterion, verification criterion, pseudo-science, empirical evidence

Other concepts used

Affirming the consequent, phrenology, non-science, logical fallacy

Themes and areas of knowledge

AOK: Natural Sciences

Recap and plan

We have agreed that bias in natural sciences is defined in relation to correspondence to reality. A belief is biased if it does not correspond to how things actually are.

But how exactly do we establish if a belief corresponds to reality? The only access to reality that we have is through experiments, but there is always a possibility that experiments themselves are flawed.

What we can do, however, is make sure that our knowledge is "true beyond a reasonable doubt". We acknowledge that we will never know *for certain* if a belief is true or not, but at least we can guarantee that we have done everything we can to ensure that it is. This guarantee is a sign of quality that science is supposed to provide.

The demarcation problem is the problem of distinguishing between science and non-science. This problem is fundamental because science provides the guarantee whereas non-science does not. This lesson will give an introduction into the demarcation problem.

Demarcation criteria

Criteria that draw a line between science and non-science are known as demarcation criteria.

KEY IDEA: The demarcation problem is the problem of telling the difference between science (which provides a guarantee that our knowledge is true beyond a reasonable doubt) and non-science (which does not). Demarcation criteria are criteria used to draw the line.

So, what is the difference between science and non-science? The question sounds really simple, but it has puzzled philosophers of science for centuries.

How can we establish the difference between science and non-science? (#Scope)

Give it a thought… I will give you a number of options (many of which are popular responses given by my students who are starting on their TOK journey):

UNIT 5 - Knowledge and understanding

Exhibition: Kamal, a navigation device 312
Story: The savior of mothers 313

5.1 - Objectivity, subjectivity and understanding 314
Lesson 1 - Subjectivity and objectivity 314
Lesson 2 - Understanding 320

5.2 - Knowledge and understanding in Natural Sciences 324
Lesson 3 - Determinism 325
Lesson 4 - Indeterminism 329
Lesson 5 - Scientific worldview 334

5.3 - Knowledge and understanding in Human Sciences 338
Lesson 6 - Reasons versus purposes 339
Lesson 7 - Verstehen 343
Lesson 8 - Intersubjectivity 347
Lesson 9 - Qualia (part 1) 351
Lesson 10 - Qualia (part 2) 355

5.4 - Knowledge and understanding in the Arts 359
Lesson 11 - Propositional and non-propositional knowledge 360
Lesson 12 - Van Gogh's Starry Night (part 1) 364
Lesson 13 - Van Gogh's Starry Night (part 2) 368
Lesson 14 - Three components of art: artist, creation, audience (part 1) 373
Lesson 15 - Three components of art: artist, creation, audience (part 2) 377
Lesson 16 - Aesthetic judgment: subjectivity and universality 381
Lesson 17 - Deep human response 385
Lesson 18 - Understanding in art 389

5.5 - Hermeneutics 393
Lesson 19 - Hermeneutics 393
Back to the exhibition 398

5.1 - Objectivity, subjectivity and understanding

In the first couple of lessons of this unit I will introduce the key concepts that will be used throughout the rest of it. As mentioned, I will argue that the concepts of subjectivity and objectivity are more complicated and multi-dimensional than they seem to be. I will separate two dimensions of objectivity and subjectivity – the ontological one and the epistemological one. By the end of the first lesson you will know lots of great concepts that will allow you to think much more deeply about what it means for something to be "objective". There will be many new words, but don't worry – we will apply them a lot in the rest of the book, so it will become your second nature to use these words casually in your daily conversations.

In the second lesson, we will discuss what it means to understand something and how it is different from knowing it. Again, these general principles will be later applied to specific areas of knowledge.

Lesson 1 - Subjectivity and objectivity

Learning outcomes

a) [Knowledge and comprehension] What are ontologically objective and ontologically subjective phenomena? What is epistemologically objective and epistemologically subjective knowledge?
b) [Understanding and application] What are the examples of subjective and objective knowledge of objectively existing and subjectively existing phenomena?
c) [Thinking in the abstract] Can we ever know objectively existing phenomena for what they really are?

Key concepts

Ontology, epistemology, ontologically objective phenomena, ontologically subjective phenomena, epistemologically objective knowledge, epistemologically subjective knowledge, noumenon, phenomenon

Other concepts used

Phenomenology

Themes and areas of knowledge

Theme: Knowledge and the knower
AOK: Natural Sciences, Human Sciences

Recap and plan

It is very common for people to believe that good knowledge has to be objective. "Subjective" has become a synonym for unreliable, unsupported and speculative. Similarly, if knowledge is objective, then it is, according to common belief, reliable, credible and trustworthy.

In this lesson I will try to show you that the relationship between subjectivity and objectivity is far more complex than it seems. We will introduce the difference between epistemological and ontological subjectivity and objectivity and look at the interplay between these two dimensions. This distinction will serve as an overarching idea that we will keep coming back to throughout this unit and even the rest of the book.

If you are someone who believes that good knowledge must be objective, I invite you to take a deep breath and read on. I don't promise to prove you wrong, but I promise to make you doubt.

Ontological and epistemological objectivity and subjectivity: definitions

The title of this section is not easy to pronounce. However, once you understand these terms, a lot of other ideas and knowledge concepts will fall into place.

As you might remember, philosophy may be broadly divided into two parts – ontology and epistemology. It is important to separate them, because mixing them up often results in confusion.

- Ontology is the study of being. It answers questions like "Does X exist?" For example: Does God exist? Is the Universe infinite?
- Epistemology is the theory of knowledge. It answers questions like "How do we know that X exists?" For example: Can existence of God be proven? How can we know if the Universe is infinite

Ontology
- Theory of being - Does X exist?
Epistemology
- Theory of knowledge - How do we know that X exists?

Is it possible to eliminate subjectivity from our knowledge of the world? (#Perspectives)

Ontologically objective phenomena are comprised of a range of phenomena that exist in the world around us. In other words, they are what we call "objectively existing reality". Ontologically objective phenomena are independent of the observer. Even when nobody is looking at them, they still objectively exist. Trees around you, the book you are reading, your brain cells and the electrical impulses in your brain – all of these are examples of ontologically objective phenomena. We will also refer to them as "objectively existing phenomena".

KEY IDEA: Ontologically objective (objectively existing) phenomena are independent of the observer. They exist even when nobody is experiencing them.

***Image 5**. Falling tree*

Imagine there is a deep forest. After a strong gust of wind, a tree falls in the middle of it with a crashing sound. There is no one around to hear that, though. The question is, if nobody heard the crashing sound, was there a crashing sound? Although there are some philosophers whose answer is no (they are known as *phenomenologists* – you can research this further if you'd like), the commonly accepted position is yes, there was a crashing sound. The falling tree produced certain vibrations in the air, and although these vibrations never reached a human ear, they did exist objectively. It was an ontologically objective phenomenon. Similarly, the forest itself, according to the common belief (but not phenomenologists!), exists even when no one is looking at it. It is just there.

Ontologically subjective phenomena are the ones that only exist in an individual's subjective experiences. You cut your finger accidentally and you feel excruciating pain – that is part of your subjective experiences. You fall in love with someone and the emotional turmoil you

UNIT 6 - Knowledge and language

Exhibition: Pioneer plaque 401
Story: Arrival 402

6.1 - What is language? 403
Lesson 1 - Signals and signs 403
Lesson 2 - Meaning 408

6.2 - Language and thought 412
Lesson 3 - Concepts 412
Lesson 4 - A priori and a posteriori concepts 417
Lesson 5 - Spacetime 421
Lesson 6 - Linguistic nativism 426
Lesson 7 - The continuity hypothesis 431
Lesson 8 - Mentalese 436
Lesson 9 - Sapir-Whorf hypothesis 441

6.3 - Language and communication 445
Lesson 10 - Translation 446
Lesson 11 - Machine translation 450
Lesson 12 - Loaded language 454

6.4 - Language in the areas of knowledge 458
Lesson 13 - The role of language in Natural Sciences 458
Lesson 14 - The role of language in Human Sciences 462
Lesson 15 - The role of language in History 466
Lesson 16 - The role of language in Mathematics 470
Lesson 17 - The role of language in the Arts 474

Back to the exhibition 478

UNIT 6 - Knowledge and language

Language is all around us. The words you are reading right now are language. The news you watched yesterday used language to tell you what happened. When you used an emoji yesterday in a text to your friend, that was language.

But more than that, language is in your thoughts. Even when you don't say anything out loud, you use language to think. Would we be capable of thinking if we could not speak a language? Some say no. Language is even in your perception. When you look at an apple and perceive an apple, you perceive an entity that you have already named. To some extent, your perception is a product of language.

Language is also a key to our culture. When a child learns a language, they internalize culture together with it. In some languages, for example, the child learns that there are two words for "you" – a polite version to be used when addressing an older person and an informal version to be used with friends. With this distinction comes the cultural attitude to old age and authority. Without language, how would culture get transmitted from generation to generation?

Language has many functions, but there are probably two key functions that everything else revolves around:

1) Language is a tool of thinking
2) Language is a tool of communication

When students think about language, they commonly assume the priority of the second function. They discuss, for example, how speaking the same language and understanding terms in the same way is important for scientists to collaborate on their work. This is indeed relevant, but I encourage you to not forget about the first function. The link between language and thought raises so many profound issues relevant to the production of knowledge. In this unit, we will consider both functions of language in turn, focusing on thinking first and then looking at communication.

Exhibition: Pioneer plaque

The Pioneer plaques are rectangular aluminum plates that were placed on board *Pioneer 10* and *Pioneer 11*, spacecraft that were launched into space in 1972 and 1973, respectively. These were the first human-built objects that escaped the Solar System. The reasoning behind the plaques was that, in case the spacecraft are ever intercepted by intelligent extraterrestrial beings (in other words, aliens), they will understand where the plaques are coming from and how to find us. We wanted to send aliens a message that they will understand even if they don't speak our language (how dare they!).

The two plaques are identical. They are 22 centimeters in width and 15 centimeters in height. Each plaque is 120 grams in weight. If you had these constraints, how would you design your message to the aliens?

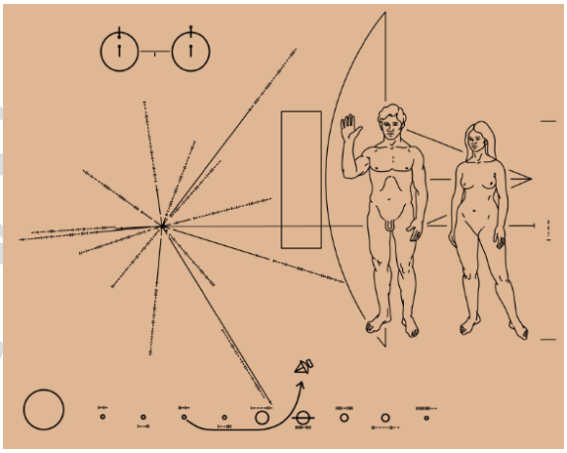


***Image 1**. The Pioneer plaque*

The figures of the man and the woman were originally intended to hold hands, but Carl Sagan (who designed the plaque) thought that aliens could misinterpret this as the man and the woman being a single creature rather than two separate creatures. The man raises his hand in a greeting gesture. Carl Sagan realized that this may not be understood by aliens, but this also shows that we have an opposable thumb and the way our arm may move.

Everything in the plaque bears significance. The radial pattern on the left, for example, shows the position of the Sun relative to 14 pulsars (pulsars are something like space lighthouses, they radiate two beams of light in opposite directions and they rotate). Most of the lines are accompanied by long binary numbers which stand for periods of these pulsars (a period is the time needed for a pulsar to make one rotation). The 15th line that extends far to the right indicates the Sun's relative distance from the center of the galaxy, using the same measurement units.

At the bottom of the plaque, there's a schematic diagram of the Solar System, also showing the trajectory of the Pioneer spacecraft travelling past Jupiter and out of the Solar System.

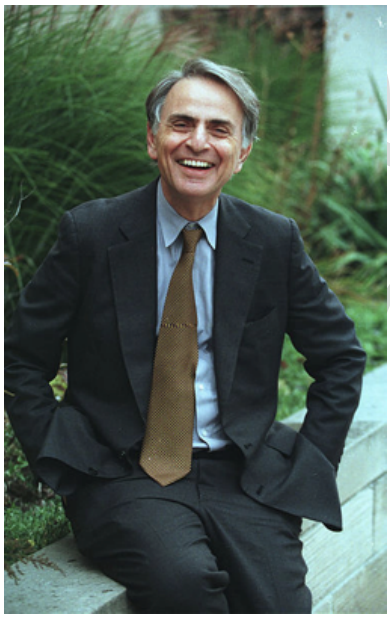


***Image 2**. Carl Sagan*
(credit: Michael Okoniewski, Wikimedia Commons)

The binary numbers near the planets show their relative distance from the Sun. The unit is 1/10 of the orbit of Mercury. The binary numbers themselves use the symbols "I" and "–" instead of "1" and "0".

Carl Sagan had only three weeks to design the plaque. Subsequently, the design was criticized for several reasons. One of them, for example, is the use of an arrow to represent the trajectory of the spacecraft. It has been claimed that arrows are so easily understood by us because we all come from hunter-gatherer societies; an alien with a different heritage may find the symbol meaningless and not suggestive of direction.

The plaques are still out there, like a message in a bottle thrown into a vast ocean. If someone finds the bottle, will they understand the message, or will they even understand that this is intended as a message? It remains an open question.

Lesson 8 - Mentalese

Learning outcomes

a) [Knowledge and comprehension] What is Mentalese (the language of thought)?
b) [Understanding and application] What evidence can support the existence of Mentalese?
c) [Thinking in the abstract] Do animals speak Mentalese?

Key concepts

Language of thought (Mentalese)

Other concepts used

Linguistic nativism, linguistic empiricism, pre-linguistic creatures, mental representations, universal grammar

Themes and areas of knowledge

Themes: Knowledge and language, Knowledge and the knower
AOK: Human Sciences

Recap and plan

We are assuming that meaning is the link between a signifier (e.g. a word) and the signified (i.e. the concept).

This raises the question of the exact nature of this link. Is it causal, that is, does one influence the other? If so, do concepts influence language or does language influence concepts? Can one exist without the other? For example, can pre-linguistic creatures have concepts?

Essentially, these are questions about the relationship between language and thought. Not easy to answer, but we will try to shed some light on it in this lesson.

What are we debating about?

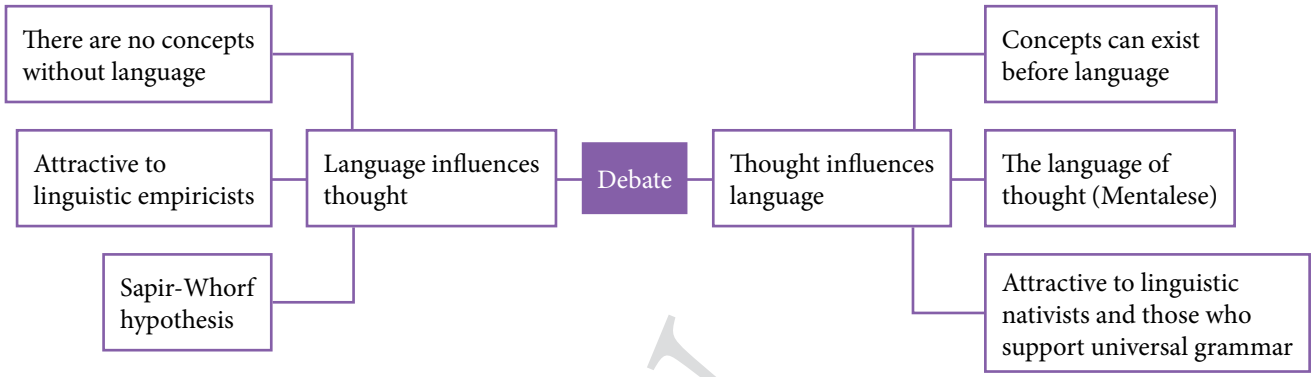
Assuming that language is secondary to thought, what implications does it have for knowledge? (#Perspectives)

The debate is about what influences what: does language influence thought or does thought influence language? I know that it is very tempting to say "both", but view this as a chicken-and-egg problem: which one is the primary influencer? Which one was there in the beginning?

***Image 32**. Language of thought*

If you belong to the camp that says "*In the beginning, there was thought, and thought influenced language*", you believe that:

1) Concepts can exist without language. They get expressed in language, but they can exist without it.
2) There is some other structure existing behind language – the "language of thought".
3) The language that we speak is an attempt to translate this "language of thought" into a conventional language accessible to others (English, Spanish, Mandarin).
4) It is possible for language to be insufficient to express the thought you want to express.



If you are with the camp that says "*In the beginning, there was language, and language influenced thought*", you believe that:

1) There are no concepts without language.
2) The "language of thought" – the hypothetical structure behind the language we speak – does not exist. This is because we think in the same language as we speak.
3) The language you speak determines the way you think and the concepts through which you understand the world.
4) The more languages one speaks, the richer their concepts and the deeper their understanding of the world.

The first position is more attractive to *linguistic nativists* and those who support the existence of *universal grammar*. The hypothetical "language of thought" that I mentioned is also referred to as Mentalese. That's the focus of the current lesson. The second position is more attractive to *linguistic empiricists* and proponents of Sapir-Whorf hypothesis (we will talk about this hypothesis in the following lessons).

KEY IDEA: Mentalese is the hypothetical "language of thought". The language that we speak is an attempt to translate Mentalese into a conventional language like English or Mandarin, but Mentalese can exist without these conventional languages.

Can there be concepts without language?

Donald Davidson, an American philosopher, thought that concepts cannot exist without language. According to him, a belief existing only as a private attitude, without being expressed in language, is "not intelligible". And therefore "a creature must be a member of a speech community if it is to have the concept of belief" (Davidson, 1975, p.170).

On the other hand, if we believe that concepts cannot exist without language, how can we explain that sometimes we have a thought that we find difficult to formulate or express? We feel like we know what we want to say, but we struggle to put it into a verbal form. Does it show that we first think (in concepts) and then speak?

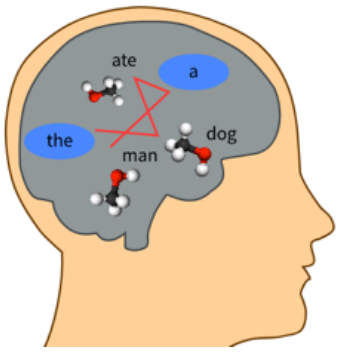


***Image 33**. Conceptual structures in the mind*

Do we think in the same language as we speak? (#Methods and tools)

Lesson 15 - The role of language in History

Learning outcomes

a) [Knowledge and comprehension] What is propaganda?
b) [Understanding and application] How can propaganda affect history writing?
c) [Thinking in the abstract] How could a historian of the future separate facts from propaganda using social media as primary evidence?

Key concepts

Propaganda, Basic English, Newspeak

Other concepts used

Fake news, propaganda bots (political bots)

Themes and areas of knowledge

Themes: Knowledge and language
AOK: History

Recap and plan

We know history through language. There is simply no other way to know events of the past. You could argue that there are videos and pictures and material evidence (such as an ancient Greek vase). But what good are these artifacts if you cannot describe what is happening in the video (using language) or explain the function of the vase (again, using language)? As you remember, history is based on the process of historical interpretation. In its turn, historical interpretation cannot exist without language.

KEY IDEA: History is based on the process of historical interpretation, but historical interpretation cannot exist without language

What is the role of language in obtaining historical knowledge? (#Scope)

There is one aspect of language use that seems to be particularly important in history – historical propaganda. This is when a historian, while describing events of the past more or less accurately, simultaneously uses language to promote his or her political agenda.

Language and propaganda

When historians create an account of events of the past, their national and political identity may make them biased. They may portray the past in a light that presents their nation, culture, political party or religious group in a more favorable light. They may also present opposing groups in less favorable light. This can be done intentionally or unintentionally. Sometimes, they may be forced to do so.

When someone describes the past with the aim of influencing opinions of others and promoting one's political agenda, this is propaganda. It is fueled by mass media and censorship.

Even if you cannot tweak the facts, you can still play with language. Seemingly, you can describe events exactly as they happened, but through the use of language you can manipulate the impressions that your audience will be left with.

***Image 57**. World War I propaganda*

Here is one example from research into the use of language in propaganda – Wegner et al. (1981). Participants in this study read one of four headlines:

1) Bob Talbert Celebrates Birthday (neutral statement)
2) Bob Talbert Is Linked with Mafia (incriminating assertion)
3) Is Bob Talbert Linked with Mafia? (question)
4) Bob Talbert Is Not Linked with Mafia (denial)

The person in question was a fictitious city council candidate several weeks before the election. After reading the headline, participants were asked to rate their impressions of the candidate. Results showed that:

What role does loaded language play in history? How is it different from other areas of knowledge? (#Methods and tools)

1) Ratings based on neutral headings (such as "Bob Talbert Celebrates Birthday") were neutral and even slightly positive. This is probably good news. It means that by default we are of a moderately positive opinion about politicians.
2) Impressions after reading the incriminating assertions (such as "Bob Talbert Is Linked with Mafia") were quite negative.
3) But, most surprisingly, ratings in the other two groups (question and denial) were as negative as ratings in the group with incriminating assertions!

What the headline says	What the audience remembers
Bob Talbert Celebrates Birthday	Okay guy
Bob Talbert Is Linked with Mafia	Bad guy!
Is Bob Talbert Linked with Mafia?	Bad guy!
Bob Talbert Is Not Linked with Mafia	Bad guy!

In other words, it does not matter if you are directly accusing a politician or simply wondering out loud if the accusation is true, or even denying it – in the mass perception the outcome will be the same: an association will be created between the politician and the accusation.

This gives you almost infinite possibilities to manipulate mass consciousness through propaganda! You are welcome. Use it responsibly.

Basic English and Newspeak

Obviously, there are also more straightforward ways to manipulate mass consciousness through language and propaganda. You can just play around with the choice of words. The same group of people may be called "freedom fighters" or "rioters" or "terrorists" depending on your perception of what they do or why they do it. Language is so rich that it has multiple labels for the same thing, each label coming with a baggage of extra connotations and associations that it triggers. Plenty to choose from!

Should historians be allowed to express personal opinions? (#Ethics)

In the early 20th century, when the world was tormented by the first global war, many scholars were concerned with the use of language in propaganda. In 1923, the English scholars K. Ogden and I.A. Richards published a book entitled *The Meaning of Meaning* where they spoke about how meaning can be abused in language. They proposed to design a new international language that would make such manipulations impossible, a language where every word has a meaning that is precisely understood by everyone. It would be a language that peels the emotive content off of words denoting facts. For Ogden, this project culminated in designing what he called Basic English – a version of the English language restricted to a core vocabulary of around 800 words, designed to convey meaning without the extra bits (a sanitized English).

UNIT 7 - Assessment guidance

7.1 - Overview of assessment in TOK 482

7.2 - TOK exhibition 483
7.2.1 - Nature of the task 483
7.2.2 - What counts as an "object"? 484
7.2.3 - TOK exhibition assessment instrument 484
7.2.4 - What should be linked to what 486
7.2.5 - Justifying the inclusion of objects in the exhibition 488
7.2.6 - Entry points 490
7.2.7 - How to structure the written commentary 492
7.2.8 - Concluding remarks 492
7.2.9 - TOK exhibition checklist 493

7.3 - TOK essay 494
7.3.1 - Nature of the task 494
7.3.2 - Typical mistakes 494
7.3.3 - Structuring the essay 497
7.3.4 - Tools of argumentation 500
7.3.5 - Communicating your ideas in a TOK essay 509
7.3.6 - TOK essay assessment instrument 514
7.3.7 - TOK essay checklist 517

some of you might have started with one or more areas of knowledge, others might have started with key overarching concepts such as "doubt" or "bias". Let's just say that you have studied some TOK "topics", and narrowing your choice of object down to one or more topics that you feel most comfortable with would be a good option. See the section "Entry points" further in this unit for a more detailed discussion of the process of selecting the objects.

2. You do not have to explain how the objects are linked to each other, and in fact they do not have to be linked

It is not an assessment requirement, and linking objects to each other will not bring you more marks. On the contrary, it may take your focus off of the more important aspects, such as the link between each individual object and the prompt.

7.2.5 - Justifying the inclusion of objects in the exhibition

Earlier I compared each object in the TOK exhibition to a sentence which together make a three-sentence answer to the question in the IA prompt. In this section I will further unpack this metaphor and clarify it.

The situation you should try to avoid is when all three objects contribute to the exhibition in the same way. In other words, you do not want them all to make the same point. Suppose you have selected the prompt "*To what extent is objectivity possible in the production or acquisition of knowledge?*" (prompt 28). Then you said something along the lines of "objectivity is impossible". You then presented the following objects:

1) a journal article arguing that global warming is not a thing,
2) a history book claiming that Columbus did not discover America,
3) a website containing a conspiracy theory that Americans never landed on the Moon.

These are interesting examples, but they all seem to make the same point: that there may always exist an alternative opinion. You have illustrated this point with the first object; the second and the third object do not seem to add anything new, they are just additional illustrations of the same idea. This is an example of when the inclusion of objects in the exhibition is poorly justified.

As the examiner is reading your commentary, you want the story to unfold before their eyes. They have understood what you are trying to say with the first object. As you move on to the second object, say something new. Obviously, it should still be clearly focused on the IA prompt. For example:

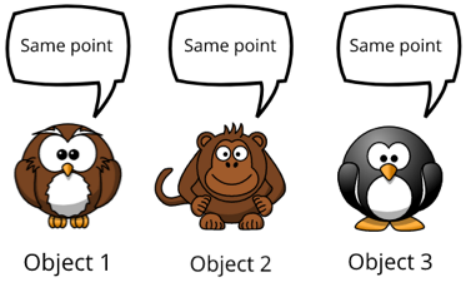


***Image 1**. Three objects supporting the same idea*

1) Your first object is a journal article claiming that global warming is not a thing. The message you are sending with the inclusion of this object in the exhibition is that there always exists an alternative opinion. Alright, that's the first sentence in your three-sentence answer.

2) For the second object you take Franz Gall's map of skull regions. In the 19th century Franz Gall, the founder of phrenology, suggested that certain cognitive abilities correspond to certain areas of the brain and that by feeling the bumps on one's skull, he

could diagnose that person's abilities. Today this is widely used as an example of pseudo-science whose claims have been debunked in multiple research studies. Today we know objectively that this opinion is not correct. Therefore, the message you are sending with this object, and the second sentence in your three-sentence answer to the question, is that some opinions are provably wrong.

3) Finally, your third object is the cover of the *Skeptic* magazine published by the Skeptics Society – a particular issue, for example, the one where they debunk Scientology (Volume 17, number 1). The existence of such societies and publications is useful to the acquisition of knowledge because, through debunking misconceptions, it lets you know which opinions are wrong and in this way contributes to your knowledge. Therefore, the claim you are making with the inclusion of the third object in the exhibition is that, although complete objectivity may not be possible, it is still possible to make progress in the acquisition of knowledge by exercising healthy skepticism.

Note that all three objects above may be related to the theme "Knowledge and the knower" – they are all about the knowledge you acquire from the stuff you read, what you choose to believe in and what you choose to dismiss.

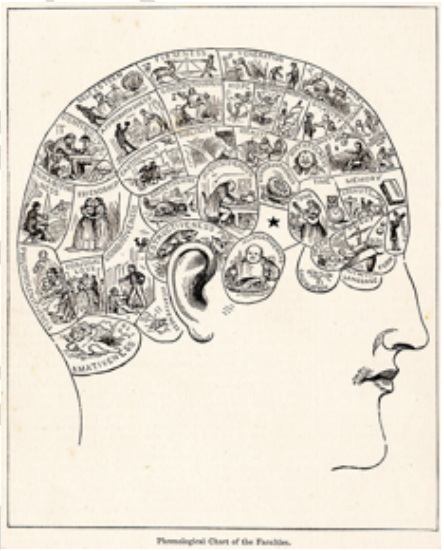


***Image 2**. Phrenology chart (1883)*

Let me just summarize. The IA prompt in this example was "*To what extent is objectivity possible in the production or acquisition of knowledge?*" The three-sentence answer that I suggested was the following:

1) There always exists an alternative opinion.
2) However, some opinions are provably wrong.
3) Although complete objectivity may be impossible, by exercising healthy skepticism we are still able to make progress in the acquisition of knowledge.

The three objects I used to illustrate these three points were a journal article that denies global warming, Gall's phrenological map of skull regions and the cover of a particular issue of the *Skeptic* magazine. Each of these objects supported one of the three points, and in this way each object made a unique contribution to the exhibition. To justify the inclusion of the object into the exhibition, then, would be to explain what unique point ("sentence") it makes in relation to the IA prompt.

I must also note that it is advisable to justify the inclusion of each object explicitly. Say "This object was included in the exhibition because…" or something similar, and then provide this justification. Examiners will be reading your commentary looking for signs of such justification. Make their job easier by simply telling them where to look.

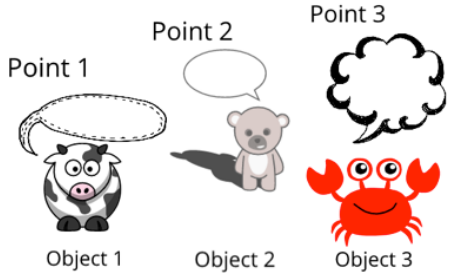


***Image 3**. Three objects each making a unique contribution*

7.2.9 - TOK exhibition checklist

In this section, you will find a checklist summarizing all of the guidance on the TOK exhibition task. You can use it to ensure that your work meets all necessary requirements. Tick the boxes that apply to your work and keep in mind your areas for improvement as you continue refining the final product.

Item	Check?
SELECTION OF OBJECTS	
I have selected one of the 35 IA prompts and I have not modified it in any way	☐
I have selected three objects, each linked to this IA prompt	☐
All of my objects are specific objects with a real-world context that I can explain	☐
Each of my three objects makes a unique contribution to the exhibition	☐
It *cannot* be said that all of my objects are examples illustrating the same point	☐
Nobody else in my class has selected the same objects	☐
COMMENTARY FOR EACH OBJECT	
I have explicitly explained the link between each of my objects and the prompt	☐
I have explained the specific real-world context of each object	☐
The real-world context of objects plays an important part in my commentary: if I remove it, the commentary will not make as much sense	☐
For each of my objects, I have explicitly formulated in one or more sentences how the object contributes to the exhibition (it has been included because…)	☐
For each of my objects, I can formulate in one sentence how it answers the IA prompt	☐
When I combine the three sentences (one for each object), they make sense as a coherent three-sentence answer to the prompt	☐
GENERIC COMMENTARY	
I have an opening statement describing my exhibition and explaining the overall message behind it	☐
I have a concluding remark that summarizes the contributions of the three individual objects and reflects on the way the exhibition as a whole answers the prompt	☐
FORMATTING	
I have included pictures of my objects in the same document with the written commentary	☐
I have included the necessary references and citations	☐
My written commentary is within 950 words	☐

IB Theory of Knowledge: A Student's Guide
By Alexey Popov

IB Theory of Knowledge (TOK) may be perplexing at first, but if you have a right approach to it, you will discover that it unlocks your critical thinking potential. This journey is full of insightful discoveries and thrilling realizations. It will remain with you when everything else you learned in school is long forgotten.

This book is designed to make your TOK journey engaging and informative. Reading this book, you will learn. You will also enhance your TOK grades, of course.

It is designed on the innovative principles of Themantic Education:

- Building blocks: important concepts are explained clearly in every lesson, with guiding questions to help you make connections and think critically.
- Thematic approach: learn all elements of the course in comparison to each other rather than separately.
- Real life: make material significant for real life and everyday situations. Not just school lessons, but life lessons.

Warning: this book contains ideas, questions and arguments that may cause you to rethink your entire knowledge and your entire life. Use it responsibly.

ISBN 9780995139008

Alexey Popov is an experienced IB teacher, author, examiner and workshop leader.