10 Statistics and probability

Mastery Professional Development

10.3 Probability

Guidance document | Key Stage 4

Click the heading to move to that page. Please note that these materials are principally for professional development purposes. Unlike a textbook scheme they are not designed to be directly lifted and used as teaching materials. The materials can support teachers to develop their subject and pedagogical knowledge and so help to improve mathematics teaching in combination with other high-quality resources, such as textbooks.

Making connections

Building on the Key Stage 3 mastery professional development materials, the NCETM has identified a set of five 'mathematical themes' within Key Stage 4 mathematics that bring together a group of 'core concepts'.

The fourth of the Key Stage 4 themes (the tenth of the themes in the suite of Secondary Mastery Materials) is *Statistics and probability*, which covers the following interconnected core concepts:

10.1 Statistical measures and analysis

10.2 Statistical representations and analysis

10.3 Probability

This guidance document breaks down core concept *10.3 Probability* into three statements of **knowledge, skills and understanding**:

10.3 Probability

- 10.3.1 Understand and use theoretical and experimental probability
- 10.3.2 Represent and calculate the probabilities of independent events
- 10.3.3 Represent and calculate the probabilities of dependent events

Then, for each of these statements of knowledge, skills and understanding we offer a set of **key ideas** to help guide teacher planning:

- 10.3.1 Understand and use theoretical and experimental probability
  - 10.3.1.1 Understand that relative frequencies tend towards theoretical probabilities as sample size increases
  - 10.3.1.2 Understand the meaning of mutually exclusive events and its implication for sums of probabilities
  - 10.3.1.3 Understand that knowledge of theoretical probabilities allows one to predict outcomes of future experiments
- 10.3.2 Represent and calculate the probabilities of independent events
  - 10.3.2.1 Broaden the representations used to identify all possible outcomes in a given situation
  - 10.3.2.2 Understand how to choose and use a representation appropriate to the given situation
  - 10.3.2.3 Understand how to calculate probabilities of independent events (and appreciate why)
- 10.3.3 Represent and calculate the probabilities of dependent events
  - 10.3.3.1 Understand the difference between dependent and independent events

10.3.3.2 Understand how the calculation of probabilities of combined events is affected by dependence/independence

10.3.3.3 Understand how to choose and use a representation appropriate to the given situation

10.3.3.4 Find and use expected frequencies from Venn diagrams, two-way tables and tree diagrams

Overview

Probability is an area of mathematics that students encounter throughout their everyday lives. This can be beneficial in providing ample opportunities to make examples tangible and relevant. It also provides challenge, as students can bring preconceptions that are grounded in instinct rather than sound logic. Awareness of this is important as students develop a more sophisticated approach to probability at Key Stage 4, as their understanding of the world will affect how readily they can identify dependence in combined events, or appreciate the effect of bias.

This core concept is particularly concerned with distinguishing between theoretical and experimental probability, and common representations used in both cases. Students should be supported to recognise that theoretical probabilities provide a model of what to expect, and that this model becomes valid when considering long-term behaviour. Emphasis is also placed on recognising that the value of the estimated probability produced by experimental data tends towards the theoretical value of the probability as the number of trials increases. This is fundamental to students developing an appreciation of the law of large numbers. As the importance of considering the number of trials is established, students can develop an awareness of the reliability of conclusions made about bias based on experimental data.

The ability to generate theoretical sample spaces for both single and combined events is developed at Key Stage 4 to include the representation of independent and dependent events. A clear distinction must be made between dependent events, that are affected by previous outcomes; and independent events, that are not. This provides an opportunity to identify and challenge the common belief (‘gambler’s fallacy’) that a certain random event is more or less likely to happen based on the outcome of a previous event or series of events. Tree diagrams can support students in identifying the different types of combined events, and provide an accessible way of analysing the probabilities of dependent events. Encourage students to make connections between the different ways of representing possible outcomes and probabilities and recognise when one representation may be more appropriate than another.

With probabilities being an example of a proportional relationship that can be expressed as a fraction, percentage or decimal, students’ understanding of probability benefits from the foundational work in proportional reasoning that takes place in Key Stages 2 and 3. When working with probabilities at Key Stage 4, it is important that teachers attend to students’ confidence and accuracy in working with fractions, decimals and percentages, so as not to inhibit the development of their understanding of probability and its associated outcomes. This is especially true of dependent events where the denominators of the fractions used to express the probabilities are affected by the outcome of prior events, and so may not always readily simplify.

Throughout the examples in this core concept, care has been taken to include a range of different data sources. Where data have been generated to exemplify a particular learning point, this often means the data have been curated in such a way that the mathematical structures are made more evident. Authentic ‘real-world’ data have also been used, so students are able to see the possible applications of the mathematics that they are learning. However, the nature of such data is they it can be more challenging to work with, and so teachers will need to consider carefully the order of examples and what their students are ready for. Where real-life data are used to generate probabilities, such as lifetime risk of a disease from the rates at which that disease is contracted, it is important to be clear about the assumptions made when modelling a question mathematically. Teachers will also need to be sensitive to the varying experiences that students will bring to the classroom, ensuring that they feel their contributions are valued, while ensuring that any misconceptions or unintentional biases are addressed.

Prior learning

Students' understanding of randomness and the ability to identify that some events are equally likely, and some aren't, is key to them being able to accurately quantify probabilities. They will have started to formalise this understanding through teaching of probability at Key Stage 3. However, it is also quite possible that their own experiences of probability in real life – for example, through games or sport – will be influential in their thinking. It is important to check students' grasp of this before further developing their understanding of probability.

Multiplicative reasoning underpins many aspects of mathematics, and probability is no exception. Reasoning in probability depends on an ability to apply proportional reasoning, involving the use of fractions, decimals and percentages. Students should already be aware that numbers can be expressed in different ways, and be comfortable converting between different forms. It is particularly important to ensure students are confident with multiplying and dividing by numbers between zero and one and to check that they know, understand and are fluent with a range of calculation strategies when adding, subtracting, multiplying and dividing fractions.

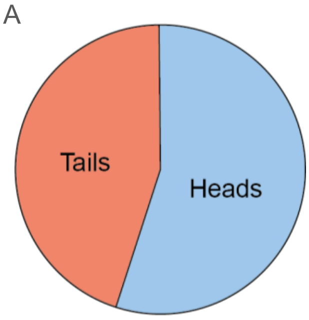
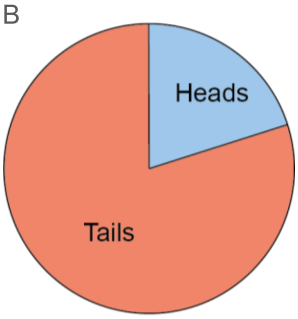
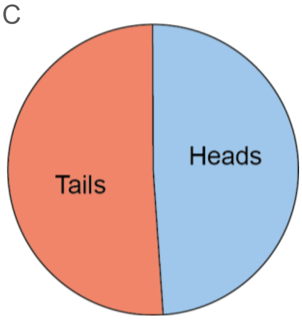
When exploring more-complex probability scenarios, students may be required to express a probability algebraically. This can involve manipulating and simplifying algebraic expressions, including algebraic fractions, or expressing chains of reasoning to prove an outcome. All these skills are covered either in Key Stage 3 or elsewhere in Key Stage 4, but students may not automatically transfer them to different areas of mathematics. Ensuring students are fluent with the use of algebraic expressions to represent situations will be useful in supporting students to determine and interpret probability algebraically.

The core concept documents 5.3 *'Probability'*, 3.1 *'Understanding multiplicative relationships'*, 2.1 *'Arithmetic procedures'* and 1.4 '*Simplifying and manipulating expressions, equations and formulae* from the Key Stage 3 PD materials all explore the prior knowledge required for this core concept in more depth.

Checking prior learning

The following activities from the NCETM Secondary Assessment materials, Checkpoints and/or Key Stage 3 PD materials offer a sample of useful ideas for assessment, which you can use in your classes to check understanding of prior learning.

Reference	Activity
Secondary Assessment materials page 47	Some counters are placed in a bag. You win if you pull a red counter out of the bag. Which of these would you choose? • Bag A has one red counter and two blue counters. • Bag B has ten red counters and twenty blue counters. • Bag C has 100 red counters and two hundred blue counters. Explain your thinking.
Key Stage 3 PD materials document 5.3 '*Probability*', Key idea 5.3.1.3, Example 4	Matilda flipped a fair coin 200 times. She drew a pie chart to show the results after 10 flips, after 50 flips and after 200 flips. Which of the pie charts (shown below) do you think shows the results after each number of flips? Explain how you know.

<table>
<tr>
<td></td>
<td>A

B

C
 </td>
</tr>
<tr>
<td>Key Stage 3 PD materials document 5.3 ‘<i>Probability</i>’, Key idea 5.3.3.1, Example 4</td>
<td>For each scenario, roll a dice 30 times and record the resulting number of points.
a) You score a point every time you roll an even number.
b) You score a point every time you roll a 6.
c) You score a point every time you roll a number less than 6.
d) You score a point every time you roll a number greater than 2.
e) You score a point every time you roll a square number.
f) You score a point every time you roll a number less than 10.
In each case, decide whether you were luckier or not than expected. Explain your reasoning by comparing your results with the expected number of occurrences.</td>
</tr>
<tr>
<td>Key Stage 3 PD materials document 5.3 ‘<i>Probability</i>’, Key idea 5.3.1.3, Example 5</td>
<td>Which of these tables do you think might be the results of flipping a coin that had been tampered with so that it is weighted to one side?
<i>Table 1</i>
<table><tr><td>Heads</td><td>3</td></tr><tr><td>Tails</td><td>5</td></tr></table>
<i>Table 3</i>
<table><tr><td>Heads</td><td>203</td></tr><tr><td>Tails</td><td>189</td></tr></table>
<i>Table 2</i>
<table><tr><td>Heads</td><td>21</td></tr><tr><td>Tails</td><td>34</td></tr></table>
<i>Table 4</i>
<table><tr><td>Heads</td><td>18</td></tr><tr><td>Tails</td><td>11</td></tr></table>
</td>
</tr>
</table>

Key vocabulary

Key terms used in Key Stage 3 materials

- combined event
- conditional probability
- dependent and independent events
- mutually exclusive events
- probability
- sample space
- Venn diagram

The NCETM’s mathematics glossary for teachers in Key Stages 1 to Key Stage 3 can be found here.

Key terms introduced in the Key Stage 4 materials

<table>
<tr><th>Term</th><th>Explanation</th></tr>
<tr><td>bias</td><td>The presence of factors that mean that the likelihood of an event occurring differs significantly from the theoretically expected value. Can also be used to refer to the unfair sampling of a population.</td></tr>
<tr><td>conditional probability</td><td>The probability of an outcome given a particular condition already being in place. For example, $P(A|B)$ is the probability of event A occurring given that event B has already happened.</td></tr>
<tr><td>expected frequency</td><td>The number of times that an outcome can be expected after a given number of trials.</td></tr>
<tr><td>frequency tree diagram</td><td>A representation that models the frequencies of all possible outcomes or categories in a given situation. The number of times each outcome has occurred is shown on its respective ‘branch’.</td></tr>
<tr><td>probability tree diagram</td><td>A representation that models the likelihood of all possible events in a given situation. The probability of each event is shown on its respective ‘branch’. For each set of branches, the probabilities should sum to 1 as they represent all mutually exclusive outcomes.</td></tr>
<tr><td>relative frequency</td><td>The frequency of a particular outcome expressed as a proportion of total outcomes. Also known as ‘experimental probability’.</td></tr>
<tr><td>set notation</td><td>A set of symbols and terminology used to denote lists of objects or outcomes and the relationships between them, often in conjunction with Venn diagrams. The list below is not exhaustive, but contains the symbols most salient for study at Key Stage 4:
<table>
<tr><th>Symbol</th><th>Term</th><th colspan="2">Example</th></tr>
<tr><td>∈</td><td>element</td><td>$x \in A$</td><td>x is an element of set A</td></tr>
<tr><td>∪</td><td>union</td><td>$x \in (A \cup B)$</td><td>x is an element of A or B or both</td></tr>
<tr><td>∩</td><td>intersection</td><td>$x \in (A \cap B)$</td><td>x is an element of both A and B</td></tr>
<tr><td>^{c} or’</td><td>complement</td><td>$x \in A'$</td><td>x is NOT an element of set A</td></tr>
<tr><td>∅</td><td>the empty set</td><td>$A = \emptyset$</td><td>there are no elements in set A.</td></tr>
<tr><td>ε</td><td>the universal set</td><td></td><td></td></tr>
</table>
</td></tr>
</table>

Knowledge, skills and understanding

Key ideas

In the following list of the key ideas for this core concept, selected key ideas are marked with a . These key ideas are expanded and exemplified in the next section – click the symbol to be taken direct to the relevant exemplifications. Within these exemplifications, we explain some of the common difficulties and misconceptions, provide examples of possible student tasks and teaching approaches and offer prompts to support professional development and collaborative planning.

10.3.1 Understand and use theoretical and experimental probability

Students work formally with probability for the first time at Key Stage 3. The introduction of simple probability experiments provides an opportunity for students to interpret the difference between how likely an event is to occur and how frequently an event occurs during an experiment.

As both theoretical and experimental probability are explored further at Key Stage 4, an emphasis can be placed on the relationship between the two types of probability and how the number of trials has an important role to play in the extent to which the experimental probability reflects the expectation based on theory. The value of experimental probability, when theoretical probabilities cannot be calculated, provides a basis for exploring bias and a means of creating a model to quantify probabilities based on relative frequency. Similarly, knowledge of theoretical probabilities allows predictions of future experiments to be made, and the results of the experiments compared and analysed to identify what the outcomes tell us about the nature of the probabilities.

Problems involving mutually-exclusive events explore students’ understanding that the probabilities of all possible outcomes sum to 1, and the calculation of theoretical probabilities contrasted with experimental outcomes reinforces the relationship between these two probability types.

10.3.1.1	Understand that relative frequencies tend towards theoretical probabilities as sample size increases
10.3.1.2	Understand the meaning of mutually-exclusive events and its implication for sums of probabilities
10.3.1.3	Understand that knowledge of theoretical probabilities allows one to predict outcomes of future experiments

10.3.2 Represent and calculate the probabilities of independent events

At Key Stage 3, students develop an understanding of a variety of ways of representing possible outcomes and their frequencies when carrying out simple probability experiments. The ability to identify and use a sample space grid is fundamental to students’ understanding of theoretical probability. The foundational work from Key Stage 3 is developed further at Key Stage 4 to include more complex situations. Tree diagrams, introduced as a way of modelling probability, provide a means of showing a sequence of events. As students’ familiarity with multiple ways of representing possible outcomes deepens, emphasis is placed on the distinctions between the different representation types and connections made between characteristics of the given situation, and the most appropriate diagram to select to represent it.

Students will have met independent events at Key Stage 3, when carrying out simple probability experiments involving coins and dice, for example. At Key Stage 4, this is formalised further and the distinction between mutually-exclusive and independent events emphasised. As students’ understanding of how to calculate probabilities deepens, their grasp of independent events can be developed. This includes an appreciation of why one event is independent of another, and the effects on the associated probability calculations.

10.3.2.1 Broaden the representations used to identify all possible outcomes in a given situation

10.3.2.2 Understand how to choose and use a representation appropriate to the given situation

10.3.2.3 Understand how to calculate probabilities of independent events (and appreciate why)

10.3.3 Represent and calculate the probabilities of dependent events

At Key Stage 3, scenarios involving bags containing beads or marbles may have been explored, and the corresponding probabilities calculated. At Key Stage 4, the introduction of dependent events provides an opportunity for successive events to be examined, where items are removed and not replaced. A comparison of experiments involving both replacement and non-replacement events can be explored to highlight the difference between dependent and independent events and the corresponding effect on the probability calculations. Expressing probabilities as fractions provides an opportunity to clearly show the distinction between the two different types of events, and to demonstrate how different combinations of events may result in the same probability, and the reasons for this.

At Key Stage 4, students' experience of probability experiments is developed to include consideration of the number of times we would expect an event to occur over a given number of trials that take place during an experiment. Having been introduced to tree diagrams as a means of representing possible outcomes and their corresponding probabilities, students can use this type of representation alongside two-way tables and Venn diagrams to find and use expected frequencies for a variety of situations. As students' familiarity with the different representations continues to develop, their understanding of when and how to use a representation that is most appropriate to the given situation can be deepened.

10.3.3.1 Understand the difference between dependent and independent events

10.3.3.2 Understand how the calculation of probabilities of combined events is affected by dependence/independence

10.3.3.3 Understand how to choose and use a representation appropriate to the given situation

10.3.3.4 Find and use expected frequencies from Venn diagrams, two way tables and tree diagrams

Exemplified key ideas

In this section, we exemplify the common difficulties and misconceptions that students might have and include elements of what teaching for mastery may look like. We provide examples of possible student tasks and teaching approaches (in italics in the left column), together with ideas and prompts to support professional development and collaborative planning (in the right column).

The thinking behind each example is made explicit, with particular attention drawn to:

Deepening	How this example might be used for **deepening** all students' understanding of the structure of the mathematics.
Language	Suggestions for how considered use of **language** can help students to understand the structure of the mathematics.
Representations	Suggestions for key **representation(s)** that support students in developing conceptual understanding as well as procedural fluency.
Variation	How **variation** in an example draws students' attention to the key ideas, helping them to appreciate the important mathematical structures and relationships.

In addition, questions and prompts that may be used to support a professional development session are included for some examples within each exemplified key idea.

These are indicated by this symbol.

10.3.1.1 Understand that relative frequencies tend towards theoretical probabilities as sample size increases

Common difficulties and misconceptions

When working with experimental probability, it is important for students to appreciate that predictions of likelihood do not predict individual events. While for an experiment with only a small number of trials it is possible for the relative frequency to be very different from the theoretical probability of the same event, it will tend towards the theoretical probability as the number of trials is increased. Students often assume that the results of an experiment will mirror the theoretical probability, regardless of the number of trials, and can misidentify the presence of bias as a result. Establishing the importance of sample size provides the opportunity for students to become more accurate with their identification of bias and to recognise why the results from a trial may conflict with expected results.

Students can find that their own experiences of certain scenarios influence their perception of likelihood and randomness. As such, it is important that they encounter a variety of contexts in the classroom. It is helpful for students to generate their own data from trials in the classroom (for example of dice rolls, coin flips or spinners), so that they can see 'live' how this tends towards theoretical probability. In a class of 30 students, if a student rolls a dice 100 times and records the data, then 3 000 trials are achieved very quickly – students can see how their own data change as they increase the number of trials, and then when all of the results for the class are collated.

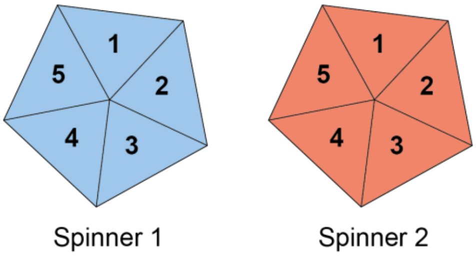
Students need to

Guidance, discussion points and prompts

Appreciate the difference between theoretical and experimental probabilities

Example 1:

Ali and Baz make two five-sided spinners, as shown below:



They decide to use the spinners to generate numbers for a bingo game. They each make their own bingo card:

6	5
7	4

Ali

2	10
3	4

Baz

Ali chooses the most likely outcomes and is confident of winning. Ali and Baz then spin the two spinners until one of them wins. They get the following totals:

6, 7, 4, 2, 6, 7, 7, 10, 6, 6, 4, 6, 3

a) Who wins the game?

b) What is wrong with Ali's thinking?

In *Example 1,* students explore expectations based on theoretical probability, alongside the corresponding results from an experiment. Genuine results are used, randomly generated here. They conflict with Ali's belief that choosing the most likely outcomes will guarantee success. This can help with **deepening** students' understanding that predictions based on likelihood may not always be reflected in experimental data.

Despite Ali including only the most likely totals and being 'in the lead' for most of the game, Baz wins the game after 13 spins of the spinners. The sample space grid to the right provides a useful **representation** to demonstrate the possible thinking behind Ali's choice of numbers. It is important to recognise that a total is not guaranteed just because it is the most probable or has a higher probability than another total.

Spinner 1

Spinner 2

	1	2	3	4	5
1	**2**	**3**	**4**	**5**	**6**
2	**3**	**4**	**5**	**6**	**7**
3	**4**	**5**	**6**	**7**	**8**
4	**5**	**6**	**7**	**8**	**9**
5	**6**	**7**	**8**	**9**	**10**

In this example, a data set has been provided, but you may choose to demonstrate randomness by repeating the scenario to create new data. Discuss with your team the benefits and challenges of asking students to carry out the experiment themselves using physical spinners, rather than producing the data via a random number generator. For example, it is important that students have opportunities to practice an organised and systematic approach when collecting and representing experimental data, but it can be time-consuming.

Understand that relative frequency can be used as an estimate of probability when theoretical probability is unknown

Example 2:

Chad and Dee are playing a board game where players need to roll an even number to advance. The only dice they can find to use has been weighted.

They want to check that the chance of getting an even number is at least 50%.

They roll the dice 200 times to check.

a) Use the table below to determine whether or not Chad and Dee should use the weighted dice.

In real-life situations, theoretical probability is sometimes unknown. Simulations can be used to generate data, or existing data can be used to determine an estimate for the required probabilities. In *Example 2*, students Chad and Dee use the relative frequency as an estimate of the probability to decide whether or not to use the dice, and an opportunity is provided to discuss the importance of sample size. Students should be encouraged to use correct **language** when describing different probability types. The terms 'experimental probability' and 'relative frequency' are often used interchangeably. Students need to understand that experimental probability is the relative frequency of an event based on collected data.

The probability of obtaining an even number using the weighted dice is $P(2) + P(4) + P(6)$. Using the relative frequencies to estimate the probabilities gives:

Number on dice	*Frequency*
1	*15*
2	*42*
3	*50*
4	*28*
5	*33*
6	*32*
Total	***200***

b) Comment on Chad's and Dee's method for finding the relative frequency.

$$P(even\ number) = \frac{42}{200} + \frac{28}{200} + \frac{32}{200} = \frac{102}{200}$$

Chad and Dee want the probability of getting an even number to be at least 50%. A probability of $\frac{102}{200} = 51\%$ satisfies their requirements.

Discuss adapting the problem to allow for the use of a computer simulation (examples can be readily found online) to explore the extent to which the results for a small number of trials differ from long-term behaviour. Consider as a team how you would manage this in the classroom, and what learning you would intend students to experience.

Example 3:

A railway company advertises their service as 99% reliable.

a) If there are 15 trains running in a day, how many trains would you expect to be on time?

b) If there are 105 trains running in a week, how many trains would you expect to be on time?

Is 99% an accurate estimate of reliability if:

c) In a month, there are 450 trains running and 14 are late?

d) In a year, there are 5475 trains running and 82 are late?

It is important that students do not solely associate the word 'trial' with stereotypical probability contexts such as dice, spinners or random selection. This example demonstrates that the principle of using relative frequency as an estimate for probability can apply in many different contexts. Here, the trials are trains arriving on time, or not. Students need to appreciate that a sample of a single day or month is not sufficient to draw conclusions about reliability. This should help with **deepening** understanding of the effect of the number of trials on the reliability of the probability estimations, which is explored further in later examples.

The **language** of expectation is first used in *Example 3* then used consistently throughout. It is important that students connect the idea of expected frequencies with the concept of probability. Students should understand that stating something has a 99% probability is the same as stating, 'We would expect this to happen 99 times out of 100'. They also need to appreciate that this expectation will not necessarily come true!

Appreciate the effects of sample size on reliability

Example 4:

Em, Fi, Gerry and Hafizah are carrying out experiments with a coin. They each flip the coin different numbers of times and record their results.

Em

Heads	5
Tails	5

Fi

Heads	7
Tails	13

Gerry

Heads	17
Tails	33

Hafizah

Heads	58
Tails	142

Example 4 focuses on **deepening** students' understanding of the importance of sample size, using the context of identifying bias from experimental results. Emphasise the different numbers of trials to develop students' understanding that the greater the number of trials, the more accurate the prediction. The variation between the four different trial results demonstrates that the probability of the coin landing on heads is tending towards 30%. (Finding the relative frequency of obtaining a head for all the trials combined is also roughly 30% *(*$\frac{87}{280}$ *= 0.3107 (4d.p.).)* Make connections with previous examples where it was recognised that results from trials for unbiased samples tend towards the theoretical probability as the sample size increases. It is important that students recognise that if the coin is flipped 200 times, the relative frequency will be much more representative than if the coin is only flipped 10 times.

Em concludes that the coin is fair because in her experiment it landed on heads 50% of the time.

Do you agree with Em? Give reasons for your answer.

How can this example be used to make effective links with understanding that the more times an experiment is carried out, the closer the relative frequency will be to the theoretical probability? How is using experimental probabilities when the theoretical probability is unknown the same/different to using relative frequencies when the theoretical probability can be determined?

Example 5:

Dean buys a packet of tomato seeds. Of the five seeds in the packet, four seeds successfully grow into seedlings.

Dean says, 'There is an 80% chance that a seed will germinate.'

a) Is Dean correct? Why or why not?

Esther also cultivates tomato plants. In April, she records the number of seeds planted in each tray, and the number that successfully germinate:

Planted	*16*	*12*	*20*	*20*	*12*	*20*
Germinated	*15*	*12*	*17*	*15*	*11*	*20*

b) Use the data to estimate the probability that a seed will germinate.

c) Which estimate is better, Dean's or the answer from part b? Why?

The following month, Esther plants 256 seeds.

d) How many would you expect to germinate?

Next year, Esther wants to have 1000 plants to sell at her village fete.

e) How many seeds should she plant?

Examples 2 and *3* can be considered a pair, as can *Examples 4* and 5. In both sets of examples, the intended learning point is the same, but the **variation** can be found in the contexts that are used. Here the focus is that the reliability of the probability estimate increases with the number of trials. Whereas *Example 4* was firmly rooted in the much-used scenario of flipping a coin, *Example 5* uses a real-life scenario to examine whether conclusions can reliably be drawn from the number of trials considered. As with the trains' reliability data of *Example 3*, students may not immediately recognise the individual seeds being grown as 'trials'.

There is **language** in this example specific to gardening that students may not be familiar with, although they should have encountered germination in science. It is important to ensure students see beyond words that they do not know and find the mathematical structures that sit beneath the example. You might find it helpful to support students to develop inference skills. Ask them to highlight any vocabulary they are unsure of and then re-read the rest of the question. What is their best guess for what that word might mean? Here, the definitions of 'germinate' and 'cultivate' are implied in the opening paragraph.

You know your students best. Will the presence of unfamiliar language be a barrier? How can you mitigate against the challenges of introducing new context-specific terminology, while ensuring students experience and gain confidence with situations where they do not already know all of the vocabulary?

Understand that empirical unbiased samples tend towards theoretical probability distributions, with increasing sample size.

Example 6:

Jack tosses a coin 10 times.

a) How many times would you expect it to land on heads?

Jack looks at his results and makes the statement shown below:

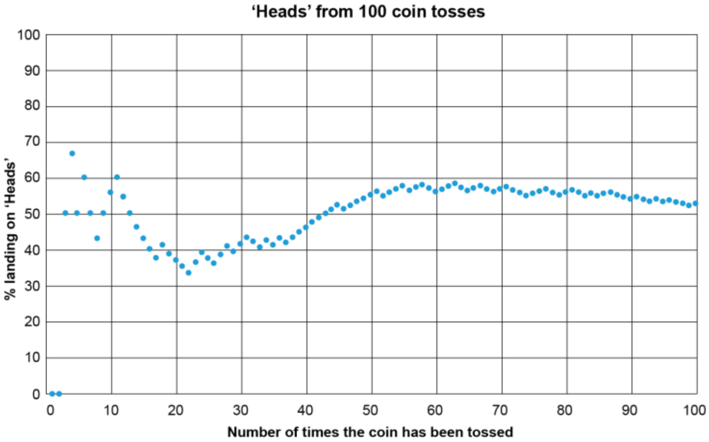
In *Example 6*, students explore the relationship between experimental and theoretical probability for different sample sizes. Part a initiates students' thinking about expectation for the results when tossing a coin 10 times. It establishes that, according to theoretical probability, the coin should land on heads for half of the tosses. Ask students to consider what result would make them think that the coin is biased towards heads. This can provide an insight into whether or not students are taking sample size into consideration. During discussions, take note of students' use of **language**. They may use 'fair' and 'unbiased' interchangeably. It is important they understand that both terms mean that all outcomes are equally likely.

I tossed a
coin 10 times and
it landed on heads
eight times. The coin
must be biased!

b) *Decide whether Jack's statement is true or false, giving reasons for your answer.*

c) *Would the coin being tossed 100 times and landing on heads 80 times be more, less or just as convincing as evidence of bias? Give reasons for your answer.*

Share a graphical **representation** of a simulation for an unbiased coin, showing the percentage of times it lands on heads plotted against the total number of tosses. This is powerful in demonstrating that as sample size increases, experimental probability gets closer to theoretical probability. The graph below has been generated using a spreadsheet. It provides a visual representation of how the experimental probability (relative frequency) of heads varies dramatically at the early stages of the experiment, then as the number of tosses increases, it begins to stabilise at around 50%. Can students predict what the graph might look like for, say, 500 tosses?



Discuss using spreadsheets to generate a similar simulation: does anyone in your team have knowledge of the relevant formulae to share? How might students use a spreadsheet to generate their own simulations? What are the benefits of this compared with carrying out a practical experiment?

10.3.2.2 Understand how to choose and use a representation appropriate to the given situation

Common difficulties and misconceptions

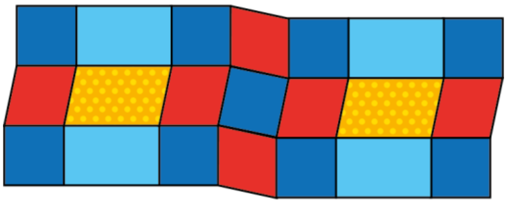
When working with probability, students are introduced to a variety of tables, grids and diagrams to help them to record and organise possible outcomes. They will often choose one representation over another, based on their affinity with it rather than an understanding of what makes representations distinct, and may select one that is not well matched to the characteristics of the data for a particular problem. Give opportunities for them to explore multiple representations for a given situation to support them in developing a deeper understanding of when a particular representation is most useful.

Students need to

Understand the mathematical structure of a Venn diagram

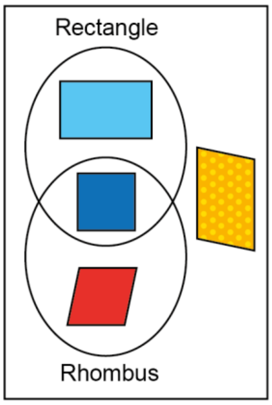
Example 1:

A tessellating pattern is made out of four different quadrilaterals:

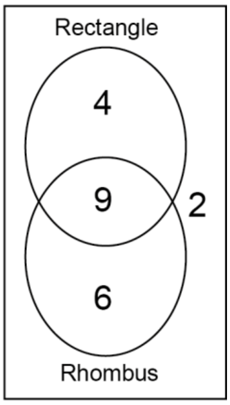


Abigail and Boki each create Venn diagrams to represent the situation:

Abigail's:



Boki's:



a) *What is the same/different about their Venn diagrams?*

b) *Use one of the Venn diagrams to complete the two-way table below. Which is more useful?*

Guidance, discussion points and prompts

Venn diagrams are a versatile **representation** when it comes to classifying according to characteristics, and one which students are likely to have experienced across the curriculum from primary school. When first introduced, students may have arranged concrete objects in the different sections of the Venn diagram based on the categories represented, before exploring pictorial representations or lists. Reminding students about sorting items into Venn diagrams can be an important step in abstracting to frequencies for work on probability at Key Stage 4. Here, Abigail's Venn diagram sorts the actual shapes whilst Boki's Venn diagram sorts the quantities of the shapes.

At this stage, the focus is on **deepening** understanding of the structure representation, rather than using the representation to find probabilities. The same context is revisited in *Example 11* for this purpose: consider how you might approach these tasks, and whether your students should experience both tasks consecutively, or would benefit more from looking at them in separate learning sequences.

Reflect on your current curriculum design. How explicitly do you teach the transition from sorting concrete items or lists to using frequencies in Venn diagrams? How might drawing attention to this shift support students to begin to find probabilities from Venn diagrams?

	Right angles	***No right angles***	***Total***
Is a rhombus			
Is not a rhombus			
Total			

Example 2:

Most artificial intelligence (AI) systems are built by people working in academia, industry, or both.

The two Venn diagrams (shown below this example) show the split of who built AI systems in 2014 and in 2024[1].

a) Comment on the differences between the Venn diagrams for 2014 and 2024.

b) How many AI systems were built with the involvement of someone from industry in 2014? How had this changed by 2024?

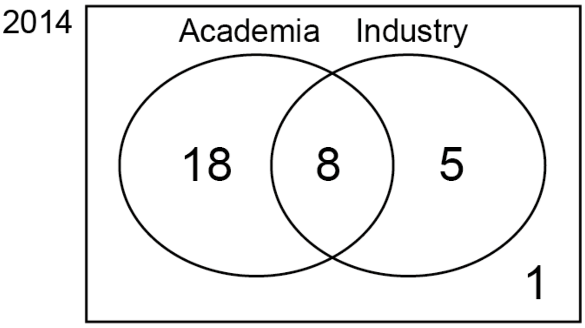
c) What proportion of AI systems were built solely by academics in 2014? How had this changed by 2024?

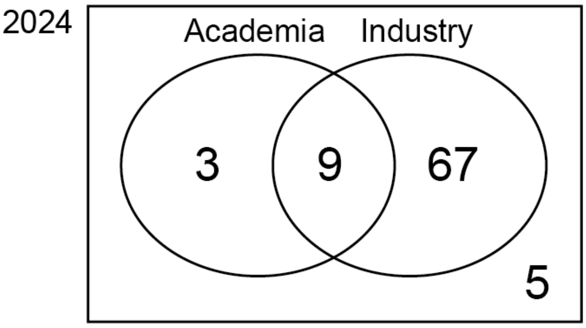
Rana says, 'Between 2014 and 2024, only one more AI system was built through collaboration between academics and industry than previously.'

d) Is Rana correct? Why does her statement not fully capture how AI design has changed in the decade between 2014 and 2024?

In *Example 2*, students are asked to compare two completely abstracted Venn diagrams that present real-world data about the origin of AI systems. The intention is that teachers can use the question prompts to explore students' understanding of the structure of the **representation**, before any further work takes place that uses the representation to explore probability.

At the time of publication, artificial intelligence is a topic that is highly relevant to students and teachers alike, and genuine data has been sourced (see references at the end of this document) for the purposes of this example. Consider the **language** in the rubric and diagrams: is it accessible to all students without further explanation? What further prompts might help students to infer the meanings of unfamiliar terms?





Example 3:

Eleanor is a dairy farmer. She has a herd of dairy cows, of which 100 give birth in spring. Information about the cows' breed (Friesian or Jersey) and number of calves born is represented in the two-way table:

	Friesian	***Jersey***	***Totals***
Single calf	*73*	*20*	***93***
Twin calves	*5*	*2*	***7***
Totals	***78***	***22***	***100***

Eleanor and her herdsman, Arthur, represent the information in two different diagrams. These are shown below this example.

a) What is the same/different about their representations?

Arthur changes the labels of the sets in his diagram. Now there are no frequencies in the overlap between the sets.

b) What might his new labels be?

c) How else could Arthur have labelled the sets of his diagram? What would change?

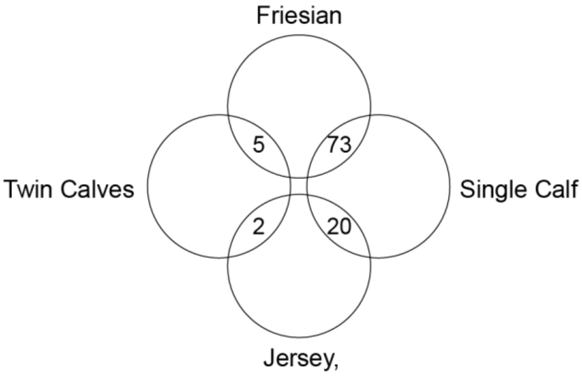
d) Use your answers to parts b and c to explain the term 'mutually exclusive'.

In *Example 3*, the comparison of Arthur's **representation** (which is a Venn diagram) with Eleanor's (which is not) makes explicit some of the features of Venn diagrams that may not be immediately obvious to students. Eleanor attempted to turn the two-way table into a Venn diagram by creating a set for every cell. Although this does show all of the information from the table, it is not a true Venn diagram as it does not show all possible combinations (which should be shown, even if they do not exist in the real world). In Arthur's Venn diagram, the sets are created by considering the two different categorisations (breed and number of calves), so that all possible combinations are now shown. This means that the two sets are labelled with one of the two options for each category, and a calf can either be in or out of each set.

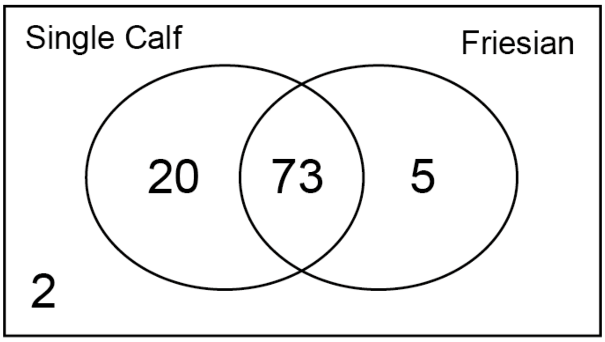
Two-way tables and Venn diagrams can offer a helpful visual to support students to refine their definitions of key probability **language** such as 'mutually exclusive'. In the case of a two-way table, mutual exclusivity is shown by the differentiation into columns and rows – a value cannot be simultaneously in both rows or both columns. In Arthur's original Venn diagram, the events are not mutually exclusive and so there are values in the overlap between the two sets. Students may, in part b, recognise that if Arthur had labelled his two sets as 'Jersey' and 'Friesian' (or 'Single' and 'Twin'), then there would be no values in the intersection as these are mutually-exclusive events. This would mean his diagram was no longer a true Venn diagram as not all possible combinations would be represented, and it would not be possible to identify the number of calves born (or the breed of the calves).

Consider whether or not to ask students to complete the three representations for themselves, rather than presenting them already completed. Discuss with your team which approach will help students to think more deeply about the characteristics of a particular representation, and how this relates to the information given. What is the minimum information you would need to provide for students to be able to do this?

Eleanor's diagram:



Arthur's diagram:



Example 4:

The legal voting age in the UK is 18. A college gathers data about the ages of students, and whether they voted in a recent general election.

	Voted	***Did not vote***	***Total***
18+	*300*	*500*	*800*
Under 18	*0*	*1200*	*1200*
Total	*300*	*1700*	*2000*

This data is arranged differently in each of the following Venn diagrams (shown below this example).

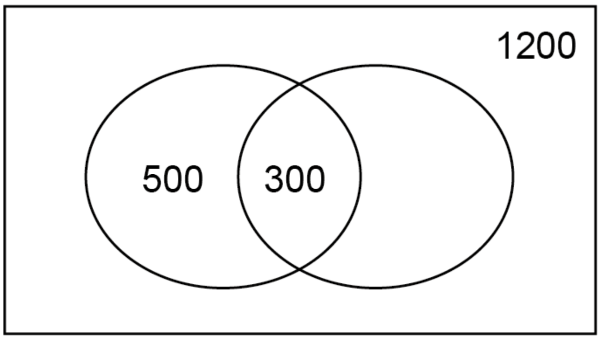
a) Label the two sets in each Venn diagram.

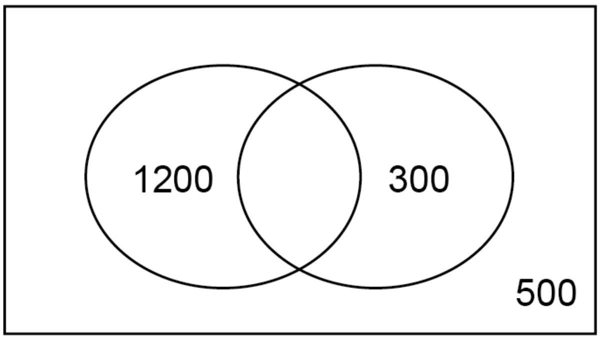
b) Which of the four Venn diagrams is the most useful representation of this data? Why?

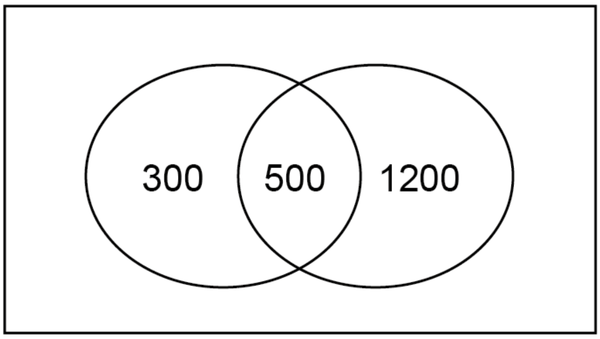
Example 4 again asks students to move between a two-way table and Venn diagram, with the **variation** designed to support them to gain deeper insight into its structure. The four different arrangements of values ensure that each value is positioned in each of the possible parts of the Venn diagram, including one of the sets, the intersection and the complement of the union.

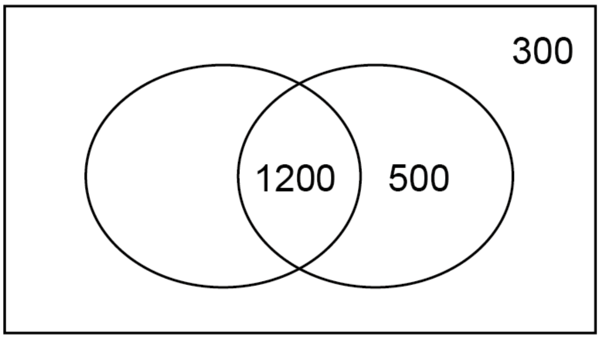
The inclusion of an empty set (the zero under 18s who voted) helps to draw attention to the relationship between sets. In the case of mutually-exclusive events – such as being under 18 and voting in a UK general election – there will be no frequencies in the intersection of two sets. For mutually-inclusive events – such as voting and being 18+ – all of the values for one of the sets appear in the intersection, leaving the rest of the set empty. Use structured **language** to support understanding of the relationships between sets: all those who voted must be 18+, but not everyone who is 18+ actually voted.

It is conventional to shade any empty sets, but at this stage in students' learning this has been omitted to ensure the focus is on the relationships depicted in the Venn diagram. Teachers might consider **deepening** students' understanding by presenting them with unlabelled Venn diagrams with one set shaded, and asking them to assign both frequencies and labels.



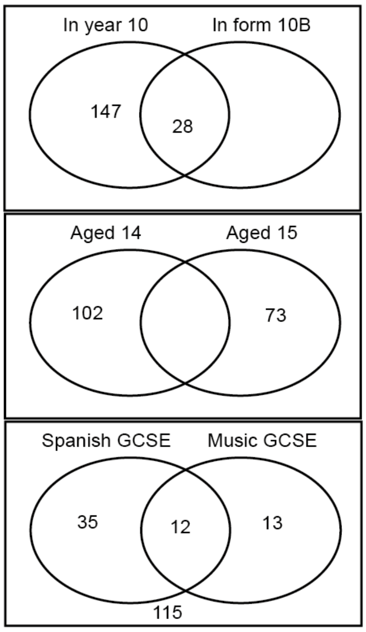






Example 5:

Vicci creates three Venn diagrams for her year group at school:

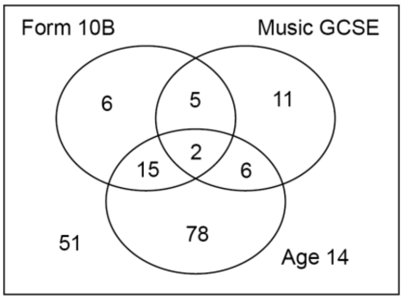


a) *Describe in words what each of the Venn diagrams shows.*

b) *How many students are in the year?*

c) *What different labels might create three similar arrangements of sets?*

Vicci makes a fourth Venn diagram:



d) *One of the values is incorrect. Which one?*

e) *Which of these questions can you now answer:*

  (i) *How many 15 year olds do Music GCSE?*

  (ii) *How many of Form 10B do Spanish GCSE?*

  (iii) *How many GCSE Spanish students are 14?*

  (iv) *How many of Form 10B are 15?*

In *Example 5*, the comparison of the three Venn diagrams offers an opportunity to discuss the features of Venn diagrams, further revealing the mathematical structures explored in the previous two examples. Reinforce the correct understanding and use of **language**, such as mutual inclusivity and exclusivity, and independence.

The **variation** of presenting the same data in different ways is intended to draw students' attention to salient features of the structure of a Venn diagram. The constant throughout – the total number of students in Year 10 – will always be found by totalling the frequencies, so each Venn diagram simply offers a different way to categorise and organise the given data.

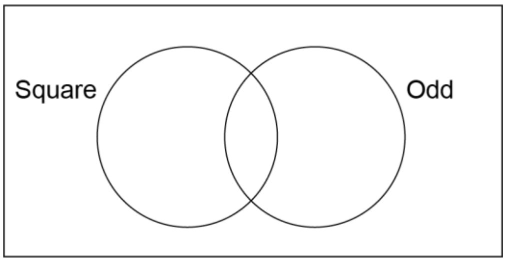
Students might reasonably argue that the first Venn diagram is not a helpful **representation** of the information. They may have seen diagrams in popular culture where one set sits entirely inside another. These are in fact Euler diagrams, mislabelled as Venn diagrams. A key feature of Venn diagrams is that all possible relationships are shown, even if they do not in reality exist – such as the impossibility of a member of Form 10B not also being in Year 10.

The final part does not ask students to find values, but to consider whether it is possible to do so. This is designed for **deepening** students' understanding about what information is made visible within a Venn diagram, and where we cannot draw conclusions without more information. Although all three Venn diagrams display data about the same 175 students, until we see the fourth Venn diagram with three sets, we do not have enough information about how the different categorisations are interrelated. As it is, we can still only answer about statements that are mutually exclusive – we can discern information about 15-year-olds, as they will not be in the 'age 14' set, but we cannot conclude anything about Spanish GCSE as some students will study Music as well as Spanish.

Begin to use set notation for Venn diagrams

Example 6:

A Venn diagram is drawn to sort the numbers 1 to 10. Set A is ‘Square’ and Set B is ‘Odd’.



a) *Complete the Venn diagram with the numbers 1 to 10.*

Below are some statements using set notation.

$$A \cup B = \{1, 3, 4, 5, 7, 9\}$$

$$A \cap B = \{1, 9\}$$

$$A' = \{2, 3, 5, 6, 7, 8, 10\}$$

$$B' = \{2, 4, 6, 8, 10\}$$

b) *What do you think the notation means?*

c) *Use your answer to part b to complete the following statements:*

  (i) $(A \cap B)' = \{\quad\}$

  (ii) $(A \cup B)' = \{\quad\}$

  (iii) $A' \cap B = \{\quad\}$

d) *Recreate the Venn diagram for the numbers 1 to 10, but with the labels ‘Prime’ for Set A and ‘Odd’ for Set B.*

e) *Complete the following statements for your new Venn diagram:*

  (i) $A \cup B = \{\quad\}$

  (ii) $A \cap B = \{\quad\}$

  (iii) $A' = \{\quad\}$

  (iv) $B' = \{\quad\}$

  (v) $(A \cap B)' = \{\quad\}$

  (vi) $(A \cup B)' = \{\quad\}$

  (vii) $A' \cap B = \{\quad\}$

A key feature of students’ maturing understanding of Venn diagrams is the introduction of set notation. This may be the first time that students have expanded their repertoire of written symbols since much earlier in their mathematics education. The introduction of any new **language** can be off-putting, and so *Example 6* returns to familiar territory: a simple Venn diagram where actual items are sorted into the sets, rather than frequencies. The intention is that students will be able to use their existing knowledge to deduce the meaning of the union, intersection and complement symbols.

The **variation** between parts a to c and parts d to e means that students can compare what is the same and what is different when only one thing changes: the ‘Square’ set changing to ‘Prime’. Prime numbers have been chosen as 2 is the only even prime, and square numbers cannot be prime, so every number in the ‘Square’ set changes, while every number in the ‘Odd’ set is rearranged. This helps draw attention to the relationships being expressed by the notation.

Example 7:

These eight statements describe different parts of the same Venn diagram either in words or set notation.

Cumin $\in A'$	*"Spice Girls who are also spices"*
Set A *= "members of Spice Girls"*	$A \cup B$ *= {Baby, Ginger, Scary, Sporty, Posh, Turmeric, Cumin, Cinnamon}*
$A \cap B$ *= {Ginger}*	*Posh* $\in A$
B' *= {Baby, Scary, Posh, Sporty}*	*'Turmeric is a member of the spices'*

a) Two statements match. Which two?

b) Write a matching statement for the remaining six statements.

c) Draw the appropriate Venn diagram.

d) Write your own statements from the Venn diagram.

Working backwards is a useful structure for **deepening** students' understanding of a concept. In *Example 7*, students are presented with summary statements and asked to create a Venn diagram for them. This helps students to practise using and building fluency with set notation, while also consolidating their knowledge of Venn diagrams and their structure.

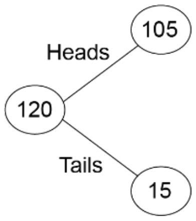
The context might amuse your staff. Do you make space for mathematically-accurate humour in your classroom? How might using a light-hearted context help make the mathematics more memorable? Discuss with your team how you might adapt this task to be more relevant and accessible to your students.

Understand the relationship between frequencies and probabilities on tree diagrams

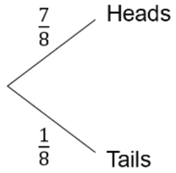
Example 8:

A biased coin is flipped, and the results represented on frequency tree and probability tree diagrams.

Frequency tree diagram:



Probability tree diagram:



When working with probability tree diagrams, students can sometimes lack a deep understanding of where the probability fractions come from and their relationship to the frequencies of outcomes. In *Example 8*, frequency tree and probability tree diagrams are presented together to help students to make connections between these two **representations**. Ensure students recognise that the starting value in a frequency tree diagram represents the total frequency (and that the numbers at the ends of arms from the same starting point add up to the starting point value) and so the coin was flipped a total of 120 times. Make connections with this being the 'whole', much like the sum of the probabilities is 1 for sets of branches.

Probability trees diagrams are often referred to as 'tree diagrams' by students, with the word 'probability' omitted. Encourage students to be accurate in their use of **language** to be able to distinguish between the two types of tree diagram, and recognise their structural differences.

Students may not recognise the relationship between the fractions in the probability tree diagram and the numbers in the frequency tree diagram just by looking at the numbers. Exploring the two diagrams in this example can provide insight into their understanding. Identifying that $\frac{15}{120} = \frac{1}{8}$ and $\frac{105}{120} = \frac{7}{8}$ is key to students understanding the relationship between the two representations. In part d, students'

a) *How many times was the coin tossed in total? Explain how you know.*

b) *How do the fractions in the probability tree diagram relate to the numbers in the frequency tree diagram?*

c) *What do you notice about the total of the fractions in the probability tree diagram? Can you explain this?*

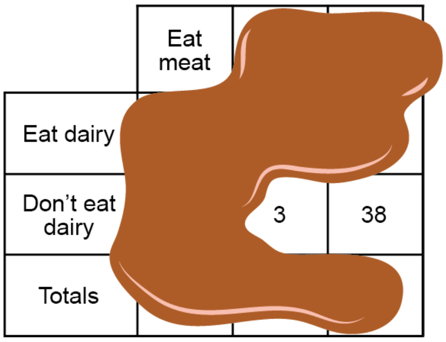
d) *If the number of times the coin landed on heads was 110, how would the frequencies and probability fractions change? What would happen to the totals?*

assumptions about the five extra heads can lead to two possible answers: they may add the extra five heads to the total frequency (so that it is 110 out of 125) or adjust the number of tails down by five (so that heads is 110 out of 120). Acknowledging the validity of both solutions provides an opportunity for **deepening** understanding of the mathematical structures. Ensure students know that, on a frequency tree diagram, the numbers at the ends of arms from the same starting point add up to the starting-point value, and this is the total frequency initially. On a probability tree diagram, the probabilities on each pair of branches always add to 1.

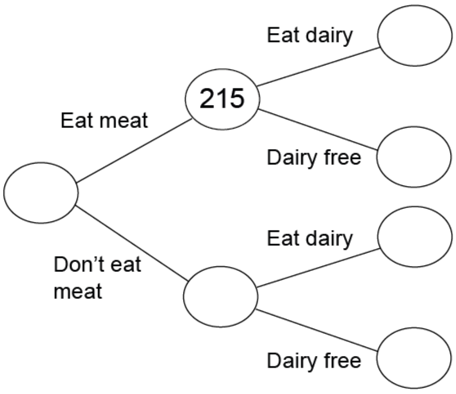
Consider using different proportions of heads and the benefits and drawbacks to using fractions that are immediately or more easily recognisable. What are the possible misconceptions about the how the frequencies combine to give the fractions? How might these misconceptions be addressed?

Example 9:

A catering company needs to plan the menu for a wedding with 250 guests. The chef spills gravy over the two-way table with information about the guests' dietary requirements:



Luckily, she had already started arranging the information into a frequency tree:



Is there enough information to complete the frequency tree with all of the guests' dietary requirements? How do you know?

Example 9 further explores the structure of a frequency tree diagram, this time extending the representation to have multiple sets of branches. Connecting the branches to a two-way table **representation** should support students to think about where each value is represented, and to appreciate that the values at the end of the two branches need to sum to the same amount as the 'root' value that they connect to.

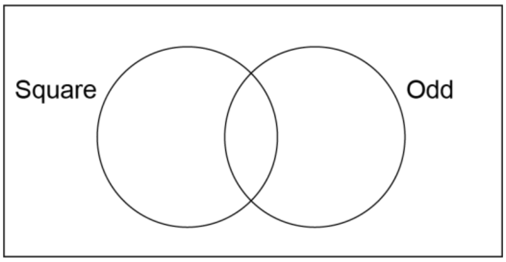
At this stage, students have not been asked to find probabilities from frequency diagrams, but this example can be used to prepare them for this next step. Drawing attention to the different 'wholes' that can be found within the frequency tree will help prepare students to identify the appropriate denominator in probability questions. Students should easily be able to read off the total of guests, meat-eaters, and those who don't eat meat. They may not initially realise that they can identify the total of guests who are dairy-free, as it requires totalling across two different sets of branches. Asking them to construct another frequency tree for the same data, but with different branches, can help with **deepening** understanding of the different ways the population can be divided.

Use different representations to calculate probabilities of independent events

Example 10:

A number between 1 and 20 is chosen at random.

Use the Venn diagram to work out the probability that the number selected is both square and odd:

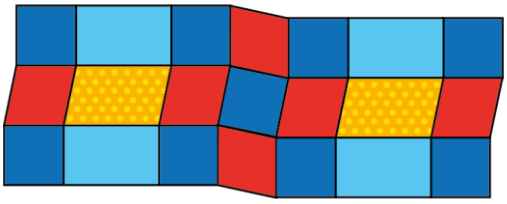


In *Example 10*, students are offered the same Venn diagram as *Example 1,* but this time need to determine the probability that a number between 1 and 20 chosen at random is both a square and an odd number. There is no direction for how students should use the **representation** provided, and so some students might position the numbers 1 to 20 in the correct sections. In this case, it is important that students are systematic when categorising the numbers to ensure that none are omitted; and that they recognise that they are looking at *the number of entries* in the intersection (2) divided by the total *number of entries* (20). Other students might categorise the values separately and then complete the appropriate parts of the Venn diagram with frequencies.

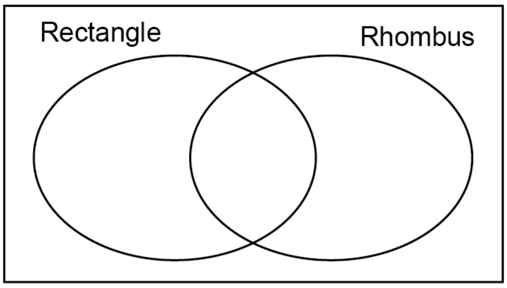
Comparison of the two approaches not only helps with **deepening** insight into mathematical structure, but also provides an interim step that students might need when determining frequencies.

Example 11:

This tessellating pattern (also seen in Example 1*) is repeated five times.*



a) *Complete the Venn diagram with the frequencies of each shape:*



The indiviual shapes are then cut up and put into a bag.One shape is chosen at random.

b) *What is the probability of selecting:*

- *(i) a rhombus*
- *(ii) a rectangle*
- *(iii) a square*
- *(iv) a rhombus without right angles*
- *(v) a quadrilateral*

Example 11 revisits the context from *Example 1* but introduces the element of chance. Students now think about how to find probabilities from the Venn diagram. The inclusion of probabilities that are 0 and 1 supports students to think about the whole set and what is/is not included. Similarly, there are some probabilities where different **language** is used to express the same thing. For example, there may be debate about whether the probability of selecting a parallelogram is 1 (as for a quadrilateral) or $\frac{4}{21}$ (as for a shape with unequal sides and no right angles).

It is likely that some students will miss the instruction that the pattern is repeated five times and complete the Venn diagram with the original frequencies from *Example 1*. This gives an opportunity to reinforce understanding of the difference between frequency and probability in Venn diagram **representations**. Whether students use the values for the 21 shapes that are visible or the 105 shapes described in the question, the probabilities will be the same.

Using a geometrical context gives an opportunity for **deepening** understanding about the overlapping properties of quadrilaterals, but also the potential for distraction from the main learning point. Students need to make connections between themes and apply their learning across concepts. Rather than avoid such tasks, instead plan carefully how they might be used in the classroom, so that any potential confusion is minimised.

When you work on this task as a team, how readily do your colleagues agree about the probabilities in part b? Is there disagreement about the definition of, say, a rectangle or a parallelogram? This might provide some insight as to how the same task *(vi) a triangle*

(vii) a parallelogram

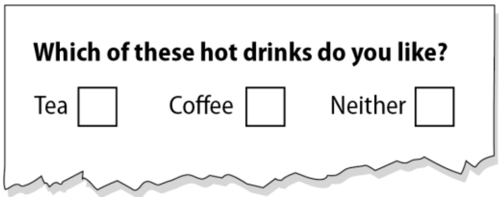
(viii) a shape with unequal sides

(ix) a shape with unequal sides and no right angles?

will be received in the classroom and help teachers prepare their responses to questions about ambiguities.

Example 12:

Kalil surveys his class about their hot drink preferences:



Which of these hot drinks do you like?

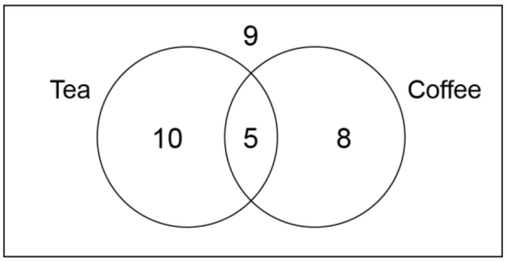
Tea ☐ Coffee ☐ Neither ☐

Kalil represents the results in a table:

Tea	*15*
Coffee	*13*
Both	*5*
Neither	*9*

Kalil wants to calculate the probability that a randomly-selected student only likes tea. He calculates the probability as $\frac{10}{42}$.

Kalil's friend Lee represents the results using a Venn diagram:



Lee calculates the probability as $\frac{10}{32}$.

a) *Explain what both Kalil and Lee have done and state who you think has completed a correct calculation, giving reasons for your answer.*

b) *What key piece of information would be useful in checking that the probability has been calculated accurately?*

Example 12 sets about **deepening** students' understanding of the potential ambiguity when it comes to the intersection of two sets. Kalil has correctly subtracted the 5 students who like both tea and coffee from the 10 students who like tea, to determine the number of students who only like tea. However, he then adds all the values to get a total of 42 for the denominator, which contains a contradiction. He has not made the same adjustment for the denominator as he did for the numerator, resulting in the occurrence of double counting. If the students who like both tea and coffee are included in the totals for those who like tea and those who like coffee (which we assume is the case as Kalil has subtracted 5 from 15 to get 10 as the numerator), then both totals need to be reduced by 5, giving a total for the denominator of 32.

Using a Venn diagram **representation** alongside a table of results shows how a Venn diagram can help to ensure that double counting of results does not occur. Lee has represented the 5 students who like both tea and coffee as the intersection of the two circles in the Venn diagram. The union is the total number of students who like tea, coffee and both tea and coffee (10 + 5 + 8 = 23) and the complement (the number of students who don't like tea or coffee) is outside of the two overlapping circles. Adding the union to the complement gives the total number of students surveyed (32). The total number of students who completed the survey provides a useful check that the Venn diagram is accurate and ensures that the table is interpreted correctly.

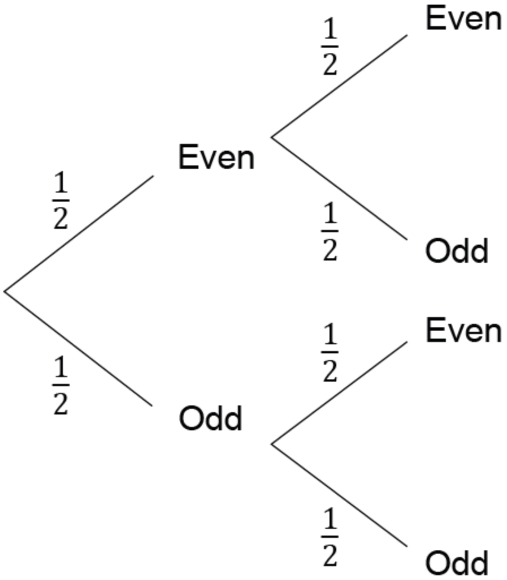
Using the **language** of 'intersection', 'union' and 'complement' with students when working with Venn diagrams can help to consolidate this new terminology. It can also reinforce understanding of the relationship between the representation and the calculations needed. This will be most relevant when determining probabilities of multiple events involving the 'and' and 'or' rules.

Data are provided for students, but carrying out a survey with the class to collect authentic data may increase student engagement. The key learning focus should still be emphasised if this is the case, so consider ways to ensure data collection is as efficient as possible. Some suggestions include surveying students as they arrive at the classroom, asking students to respond to a survey as part of the register process, or providing template tables for students to complete.

Example 13:

In a game, two dice are thrown and their totals multiplied together. Players score a point if the answer is odd.

The possible outcomes are shown on the probability tree below:



Enrico says, 'The answer can either be odd or even, so the chances of scoring are 50/50.'

a) *Use the probability tree diagram to show why Enrico is wrong.*

The possible answers are shown in the sample space diagram below.

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	8	10	12
3	3	6	9	12	15	18
4	4	8	12	16	20	24
5	5	10	15	20	25	30
6	6	12	18	24	30	36

b) *Choose a representation to explain explain why the chance of winning is 25 per cent.*

Example 13 represents the first time that these materials explore finding probabilities of combined events, and so the context is deliberately simple. Students often mistakenly think that, if there are two possible outcomes, then the probability of each outcome must be equal. A sample space **representation**, with shading to reinforce the different overall outcomes, is used as an alternative visual. It will help students to understand why the probability of an odd answer is lower than the probability of an even one – there are, simply, fewer odds than evens.

Pay careful attention to the **language** that students use in their explanations for parts a and b. They need to be confident in working interchangeably between fractions, percentages and decimals so that this does not present a barrier to reasoning with probabilities. It may also be helpful for students to organise their thinking using probability notation. For example, Enrico should have identified that there are three different combinations of dice that result in an even score, which he could have organised as:

$$P(E) = P(E,E) + P(E,O) + P(O,E)$$

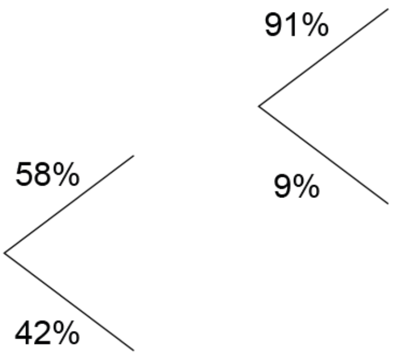
Probability notation is consistent with function notation and that is not a coincidence! Probability, $P(X)$, is the function of an event (X) that maps it to a number between 0 and 1, according to the likelihood that an event will occur. For mutually-exclusive events, $P(X)$ can be calculated as the number of ways the event can occur divided by the total number of events: in this case $P(Odd) = \frac{9}{36}$. Discuss this with your colleagues. Is the link between probability and functions something that you had already recognised? How might making this connection explicit to students support them to have a deeper understanding of functions, probability, or both?

Example 14:

The World Health Organisation (WHO) is interested in how adolescent health is affected by technology. The two-way table shows the results of their research into the gaming habits of 15-year-olds, divided by gender[2]*:*

	Girl	***Boy***
Non-gamer	*42%*	*11%*
Non-problematic gamer	*53%*	*74%*
At risk of problematic gaming	*5%*	*15%*

Some of the data are used to create a probability tree diagram:



a) What is represented by the probability tree?

b) Why are the values in the second set of branches not in the table?

c) Use the probability tree to show that the probability of a random 15-year-old girl being a gamer with unproblematic habits is 53%.

Chiamaka says, ‘I thought probabilities had to add to 100%, but 42% + 58% + 91% + 9% = 200%.’

d) What has Chiamaka done wrong?

e) Use the two-way table to create another probability tree. Is there more than one possible tree you could make from these data?

Example 14 uses real-life data, and a context that students should find interesting. The intention here is to support students to understand which probabilities need to sum to 1, so **deepening** understanding of the probability tree. Students should be able to identify which column is being represented, via the 42% of girls who are non-gamers. They may find it more challenging to understand why the branches of the second part of the tree do not just reflect the second and third rows in the two-way table. Support them to notice that 91% and 9% sum to 100% and to speculate as to why that is. Teachers might like to ask students how the diagram would look different if, instead, each event had three branches (to represent the three categories), reinforcing that this is valid so long as these three probabilities still sum to one.

In part c, students need to multiply along the branches to find the probability of being a non-problematic gamer. Students often remember rules about adding or multiplying probabilities and might be looking to see the words ‘and’ or ‘or’ in the question rubric. If it helps them to think in this way, encourage them to rephrase the scenario using that familiar **language**. They need to find the probability that the randomly-selected girl is both a gamer and at risk of problematic gaming.

Part d offers a further opportunity to clarify misconceptions about which probabilities sum to 1: students need to be clear that mutual exclusivity is the key condition here, and that each set of branches is a **representation** of all of the mutually-exclusive outcomes of a single event. In effect, the ’girl’ column of the two-way table has been split into two different events: the first is being a gamer (or not), and the second is being at risk of problematic gaming (or not).

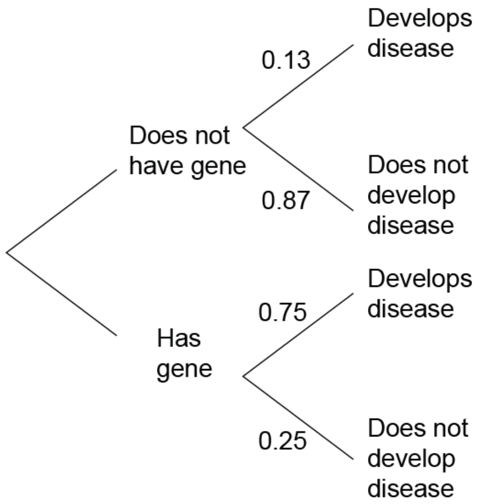
This is an example of real-world data where the information has been categorised according to gender. Students may have questions about how ‘gender’ has been defined in this survey, and teachers can refer to the source material referenced at the end of this document. Researchers need to be sensitive to the population that they are collecting data from, whilst ensuring that relevant information is collected. It will depend on the purpose of the research as to whether it is necessary to specify natal gender (for example if collecting data about health conditions) or gender identity. If teachers choose to collect data from their students, they should be aware that some categorisations may cause more issues than others, and consider the most appropriate ways to subdivide their students, being sensitive to their own context.

Example 15:

The lifetime chance of getting a disease is 13%. If a certain gene is inherited, this increases to 75%.

Pat knows that one of her parents carries the gene. Her chances of inheriting it are 50%.

a) Complete the probability tree below for Pat's situation.



b) What is the probability that Pat will develop the disease?

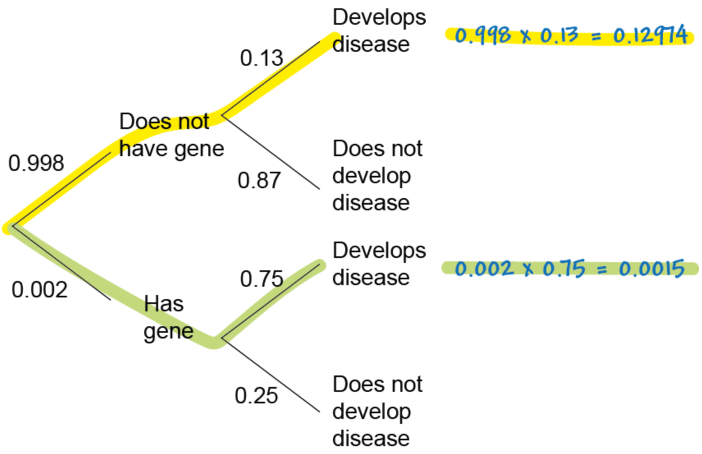
In the general population, the chances of inheriting the gene are 0.2%.

c) Create a new probability tree to show the situation for a randomly-selected person.

d) What is the probability of a randomly-selected person developing the disease in their lifetime?

This may be the first time that students have seen a probability tree **representation**, where two pairs of branches have the same outcomes but different probabilities assigned to those outcomes. Think carefully about how you will expect students to notate their workings and solutions to parts b and d, so that they can keep track of their thinking and not get muddled with the different values.

Modelling mathematical **language**, including the accurate use of notation, is key when students are starting to look at combined events where there are multiple ways to achieve the outcome being investigated. For both parts b and d, there are two 'routes' to developing the disease, and it can really help students to both visualise and organise their thoughts if these routes are highlighted, as shown below.



From here, students might record their workings as:

P(disease) = P(gene ∩ disease) + P(no gene ∩ disease)

Medicine is a field that relies heavily on probability, and produces some hugely-interesting statistics. The probabilities given here reflect a real-life gene, BRCA-2, but this is not identified in the question so as not to cause distress for any students or teachers who have personal experience of this. It is important that students can root their probability learning in real-life applications, but equally important that teachers are equipped to deal sensitively with any issues that arise from this. Discuss with your team whether they feel confident in using such data, and what support they might need.

Example 16:

For their Key Stage 4 GCSE options, students need to choose exactly one technology subject and one language. The options for technology are either Food or Product Design. The options for languages are German or Spanish.

Below this example are three different representations of the choices they made. All of them are incomplete.

a) Complete all the missing numbers.

b) Find the probability that a randomly-selected student studied:

  (i) Food GCSE

  (ii) Spanish GCSE

  (iii) Food and Spanish GCSEs.

c) Find the probability that a randomly-selected:

  (i) GCSE German student also studied GCSE Product Design

  (ii) GCSE Product Design student also studied GCSE German

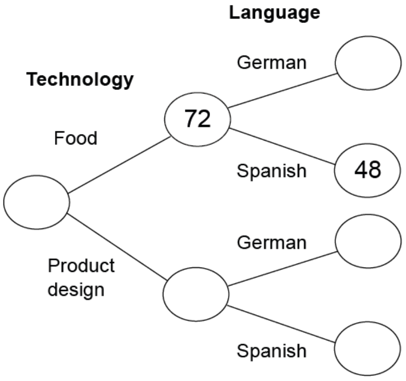
  (iii) GCSE Spanish student also studied GCSE Food

  (iv) GCSE Food student also studied GCSE Spanish.

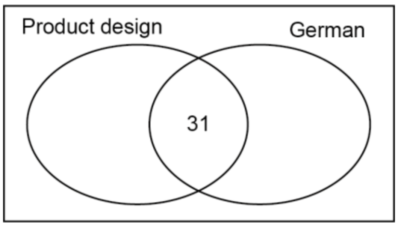
d) Which representations did you find most useful for each question in parts b and c?

The final example in this sequence brings together all of the different **representations** and asks students to work between them. Students may be surprised that there is sufficient information given, but it is possible to complete all missing values across all three representations. There is opportunity here to compare how each representation organises the information: students could reflect upon which is most useful for categorising the data, and which for calculating the given probabilities. Students need to consider whether any representation is useful for finding all of the probabilities, or if different representations are more useful for particular lines of enquiry.

Students will need to pay careful attention to the **language** in each part-question, as the differences between them are subtle. Asking students what is the same and what is different between sub-parts i and ii of part c, for example, will help them to start to think about the different possible populations they are drawing from. This is intended to consolidate work on independent events and to lay the foundations for subsequent work on conditional probability, which is explored in more detail later in this document.



	Food	PD	Total
Ger			
Span			65
Total			



10.3.3.2 Understand how the calculation of probabilities of combined events is affected by dependence/independence

Common difficulties and misconceptions
When working with probabilities, students are often presented only with problems involving fractions that can be simplified easily. As a result, they can form an implicit understanding that probabilities will always take this form. When working with a problem that involves fractions that cannot be simplified, students might assume that an error has been made. This becomes particularly relevant when working with dependent events where initial fractions (which may readily be simplified) often become more complex (and not readily simplified) due to a lack of replacement. As a result, students are often more prone to assume independence when calculating probabilities of combined events, to make the calculations easier. Presenting problems involving a range of fractions, including those that can and cannot be simplified, that are not for both dependent and independent events, can increase students' confidence in working with fractions when calculating probabilities. An understanding that probability provides a model for real-life data, which rarely results in 'nice' fractions that can be simplified, is also key to students developing an understanding of how maths relates to the real world.

Students need to	Guidance, discussion points and prompts
Know how to distinguish between independent and dependent events *Example 1:* *Which of the following scenarios are independent and which are dependent events?* *A. The outcomes of the first and second flips, when a coin is flipped twice.* *B. The probability that the cards dealt to players are spades, when one card from a 52-card pack is dealt to each of two players.* *C. The chance of rain one day and the chance of rain the next day.* *D. The likelihood of player 1 and player 2 both rolling a six to start a board game.* *E. The chance of a bus being late two days in a row.* *F. The probability of two consecutive traffic lights being red.* *G: The likelihood of a football team missing two consecutive penalties.*	In *Example 1,* students categorise scenarios based on whether the events are independent or dependent. The scenarios described expose possible misconceptions, **deepening** understanding of what it means for one outcome to be affected by another. Students may bring assumptions to the scenarios described – such as, once someone has rolled a six in a board game, it makes it less likely for the next person to roll six too. Ask students to come up with examples of independent and dependent events for themselves. They may come up with trivial independent events. While these may have been identified correctly, it is important that the focus is more nuanced, and **language** precise, to address scenarios that often result in incorrect assumptions being made. In some of the examples, classifying the event as dependent or independent is not straightforward. For example, due to the nature of weather systems, there may be occasions when the chance of rain one day is dependent on whether or not it rained the previous day. There are many factors that may need to be taken into consideration, but technically speaking, the chance of rain one day and the next are independent events. Discuss with your team: are teachers prepared to handle this ambiguity in the classroom?

Understand the effects on probabilities of replacement/non-replacement

Example 2:

A bag contains 9 sweets; 5 are hard-boiled sweets and 4 are chewy sweets. A sweet is taken out of the bag and eaten. A second sweet is then taken out of the bag.

Amira and Salote each draw a probability tree diagram, shown below this example.

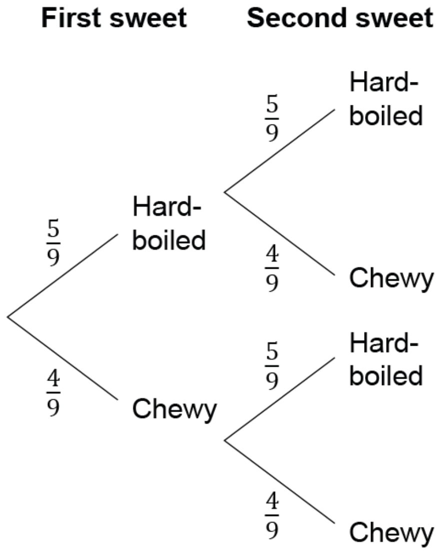
- *a) Whose probability tree diagram correctly represents the situation?*
- *b) Use the correct tree diagram to show that the probability of the first sweet being hard boiled and the second sweet being chewy is the same as the probability that the first sweet is chewy and the second sweet is hard boiled.*

The tree diagram **representation** is particularly useful when considering successive events. In *Example 2*, effects on probabilities of situations involving non-replacement versus replacement are explored. In the first tree diagram, the probabilities for selecting a hard-boiled sweet and a chewy sweet are repeated on the second set of branches due to Amira assuming the sweet is replaced. The second, correct, tree diagram, accounts for the first sweet being eaten and so the probabilities on the second set of branches differ to those on the first set.

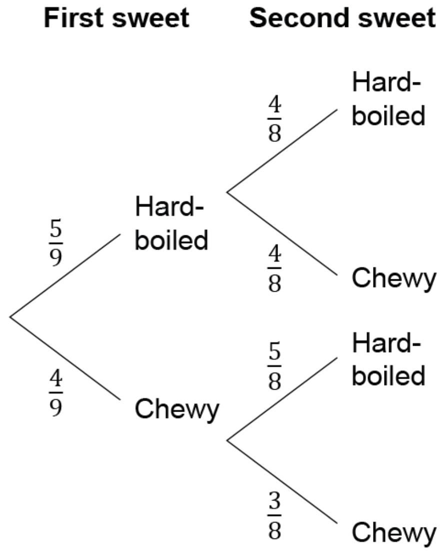
Leaving all fractions unsimplified means that the connections between the numerators and the number of each type of sweet, and the denominators and the total number of sweets in the bag can be emphasised. This helps with **deepening** students' understanding of the mathematical structure of dependent events.

Extend the discussion of why the probabilities are the same to the general case. How might this further deepen students' understanding of mathematical structure and how the tree diagram representation shows this?

Amira's tree diagram:



Salote's tree diagram



Example 3:

Bethany is calculating probabilities from a deck of playing cards. There are 52 cards in a deck, organised by colour and suit:

		Red	
		Hearts ♥	*Diamonds* ♦
Black	*Spades* ♠	*13*	*13*
	Clubs ♣	*13*	*13*

What might each of these probabilities be?

a) $\frac{1}{2}$ *b)* $\frac{1}{4}$ *c)* $\frac{1}{13}$

d) $\frac{1}{13} \times \frac{3}{51}$ *e)* $\frac{1}{4} \times \frac{1}{4}$ *f)* $\frac{1}{4} \times \frac{12}{52}$

g) $\frac{1}{2} \times \frac{25}{51}$ *h)* $\frac{1}{13} \times \frac{3}{52}$ *i)* $\frac{1}{52} \times \frac{1}{52}$

Working backwards to suggest possible questions from given solutions is a useful pedagogical structure for **deepening** understanding. In *Example 3,* the intention is that students can use their (relatively straightforward) answers to parts a to c to inform their answers to the more complex questions from part d onwards. There are multiple right answers: for example, part b could be any suit, and part c any card value. Students should be encouraged to share and compare answers, justifying which are valid. They could be challenged to think of an answer that no-one else in the class has suggested.

The **variation** in this example is designed so that teachers can compare specific pairs of questions, drawing attention to what is the same and what is different. For example, parts e and f both begin with $\frac{1}{4}$ (suggesting the selection of a particular suit) but the second fraction is the same for part e (suggesting that it is an independent event, and the card is replaced) but different for part f (suggesting it is dependent, and the card is not replaced).

In this context, students may bring a variety of different levels of experience with playing cards. A two-way table is used as a **representation** of the structure of the deck, and teachers will need to decide whether to spend time exploring this before tackling the questions, or to reference back to the two-way table as needed.

Example 4:

In a school, a Year 8 student is picked at random to be 'student receptionist' for the day. Once they have been selected, they cannot take the role again until every student has had a turn.

There are 120 students in Year 8.

a) What is the probability of being the first student to be picked at the start of the autumn term?

By the end of autumn term, there have been 75 school days.

b) What is the probability of:

(i) not yet being picked?

(ii) being the first student to be picked at the start of the spring term?

By the end of the spring term, there have been 135 school days.

c) What is the probability of:

(i) having been picked twice?

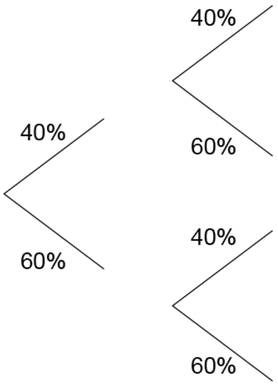
Example 4 again explores the effect of non-replacement. The **variation** is intended to allow teachers to explore different aspects of students' understanding of the relevant values and how to use these to generate probabilities. In parts a and b, the focus is on how the numerator and denominator are affected by the change in circumstances. This builds in complexity so that, by parts c and d, students are not only identifying the relevant fractions but also multiplying to combine the probabilities.

This example is offered without a **representation**; consider whether one would be helpful for students to access the question. Reflect on what aspect of the problem students are likely to struggle with, and select representations accordingly. For example, a grid with 120 squares for part a, with 75 of the squares coloured for part b, may help students to visualise the situation. This could be used in conjunction with a probability tree on a whiteboard, so that the probabilities can be erased and changed as the number of available students decreases.

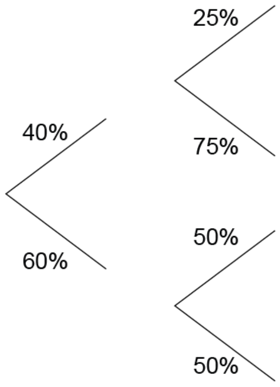
The context has been borrowed from a real school activity so it should be accessible and relevant to students. Every school has its own traditions and routines – are there any similar examples from your own setting? You could adapt this question to incorporate any

(ii) being the first student to be picked at the start of the summer term? *By the end of the summer term, there have been 190 school days.* *d) What is the probability of only being student receptionist once over the academic year?*	random selection process that your students will be familiar with.
Calculate the probabilities of combined events *Example 5:* *There are five counters, numbered 1 to 5, in a bag.* *Two different scenarios for selecting two counters are shown below this example.* *a) What is the same and what is different?* *b) What labels could go on the tree diagrams?* *Bev decides to label the branches 'prime' and 'square'.* *c) In which situation is it more likely to pick two counters that are both prime?* *d) In which situation is it more likely to pick one prime and one square?* *e) Would the probabilities be the same if Bev's counters were numbered 6 to 10? Why or why not?*	Working backwards from a **representation** offers an opportunity to make sense of the mathematical structures that are being explored. Teachers need to be aware that there are multiple possible answers to part b, and so be ready to quickly ascertain whether students' suggestions are correct. Students are likely to find it both helpful and encouraging to see their teachers 'thinking out loud' when deciding if a suggested set of labels is viable, and so teachers should not shy away from modelling their thought process when they do this. The **variation** in this example is designed to draw attention to the different number of counters in the bag in the second example. Students first need to identify the differences between the two probability trees, and recognise that the percentages in scenario 2 imply that there is something different about the counters in the bag. it is not possible to have 75% of 5 counters unless counters have been cut up. Therefore, the total of counters must have changed – the first counter selected is not replaced. If students find this challenging, converting the percentages to fractions may help – highlight the different denominators. There are lots of ways to continue and extend this task to support with **deepening** students' understanding. For example, what would be the same and different about this question if there were 10 counters originally? How about 20? In the latter case, we cannot be sure that non-replacement has happened in scenario 2, because all of the percentages of 20 result in integer solutions – can students identify this?

Scenario 1

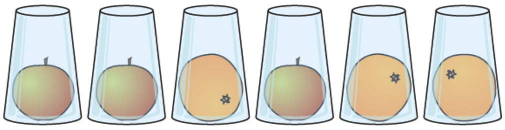


Scenario 2



Example 6:

Meg and Nel are playing a game in which three apples and three oranges are hidden under six cups (one piece of fruit under each cup):



The order of the six cups is shuffled, and two cups are picked up to reveal the fruits underneath. If the two fruits are the same, Meg scores a point. If the two fruits are different, Nel scores a point.

Is the game fair? Explain your answer.

Example 6 considers combined events in the context of the design of a game. Students may assume that, because there are two types of fruit, there is an equal chance that the two revealed fruits are the same or different, and so the game is fair. Encourage them to consider the different possible outcomes to challenge that assumption – you may need to encourage them to use **representations** to organise their thoughts. They may, for example, label the three apples A_1, A_2, and A_3 and the three oranges O_1, O_2, and O_3 and list the possible outcomes:

Same colour: A_1A_2, A_1A_3, A_2A_3, O_1O_2, O_1O_3, O_2O_3.

Different colours: A_1O_1, A_1O_2, A_1O_3, A_2O_1, A_2O_2, A_2O_3, A_2O_1, A_2O_2, A_2O_3.

Alternatively, they may represent the possible outcomes using a sample space grid.

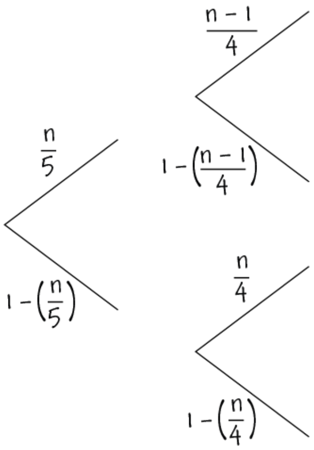
This task could be extended in many different ways, **deepening** students' understanding of the calculation of probabilities for combined events. For example, discuss whether or not the order matters (for example, is A_1A_2 the same as A_2A_1). You could also ask students to consider the design of a similar game, where the probability that both fruits are the same is the same as the probability of getting two fruits that are different. Playing the game multiple times to see the effect on the points scored compared to the one given in the example may also make for an interesting discussion.

Form and solve equations to find probabilities

Example 7:

Charlie has five school shirts. n of his shirts have a school logo on.

He usually wears a clean shirt every day, and so he represents his choice of shirt on two consecutive days using the probability tree below.



Example 7 demonstrates an important development in students' mathematics journey – recognising that algebraic terms can be used in lieu of numerical values. The Key Stage 3 and Key Stage 4 PD materials are both organised in such a way as to combine number and algebra topics, to reinforce the fact that they both rely on the same mathematical structures. At Key Stage 4, it is important that this application of algebra extends beyond the number-based themes, so **deepening** students' experience. Students' fluency with algebraic manipulation is key: in the case of working with probabilities, this includes using fraction notation to represent division with unknowns and simplifying algebraic fractions.

Teachers should be aware of the **language** used in explanations, particularly if there is a difference once algebra is introduced. Support students to make connections and feel confident by mirroring the language originally used to introduce probability tree diagrams and dependent events.

What other examples are there within your scheme of work where students can make connections between seemingly distinct strands of mathematics? Do you make the most of opportunities to introduce algebraic terms in core concepts? A useful professional development exercise might be to collaborate

a) *Add labels to the branches of the probability tree.*

b) *Show that the probability of Charlie choosing two consecutive shirts with school logos is* $(\frac{n^2-n}{20})$.

Charlie accidentally puts Monday's shirt back in the wardrobe without washing it. He picks Tuesday's shirt at random.

c) *How does this change the probability of having two consecutive logo shirts?*

d) *Draw a new probability tree to reflect this.*

The actual probability of having two consecutive non-logo shirts is 0.3 (assuming Charlie ***does*** *remember to wear a new shirt on the second day).*

e) *How many shirts does Charlie have with logos on?*

in designing some different tasks where students are required to form, solve or manipulate equations in different mathematical contexts.

Example 8:

One hundred people are surveyed: 71 are right-handed; 44 are either right-handed or have brown eyes, but not both; and 21 have blue eyes.

a) *Complete the totals in the two-way table below.*

	Right-handed	Left-handed	**Totals**
Brown eyes	a	b	
Blue eyes	c	d	
Totals			**100**

b) *Write and solve equations involving* a, b, c *and* d *to determine the probability that a person chosen at random is right-handed and has blue eyes.*

Example 8 uses the two-way table **representation** in a way that supports students in thinking deeply about mathematical structure. There is variation in how the information needed to complete a two-way table can be given. In this example, students are required to move beyond simple calculations to forming and solving equations. Once the totals have been completed, recognising that $b + c = 44$ is key to being able to determine the values of a, b, c and d (and in particular c, the number of people who are right-handed and have blue eyes).

$a + b = 79$ (1)
$a + c = 71$ (2)
$c + d = 21$ (3)
$b + d = 29$ (4)
$b + c = 44$ (5)

(1) – (2) gives $b - c = 8$ (6)
(5) + (6) gives $2b = 52$, which can be solved to give $b = 26$.

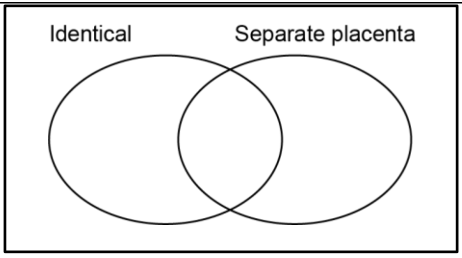
Substituting back into the equation (5) gives $c = 18$ and so the probability that a person chosen at random is right-handed and has blue eyes is $\frac{18}{100} = 0.18$.

Other than for completeness, what can identifying the values of a and d add to students' understanding of mathematical structure?

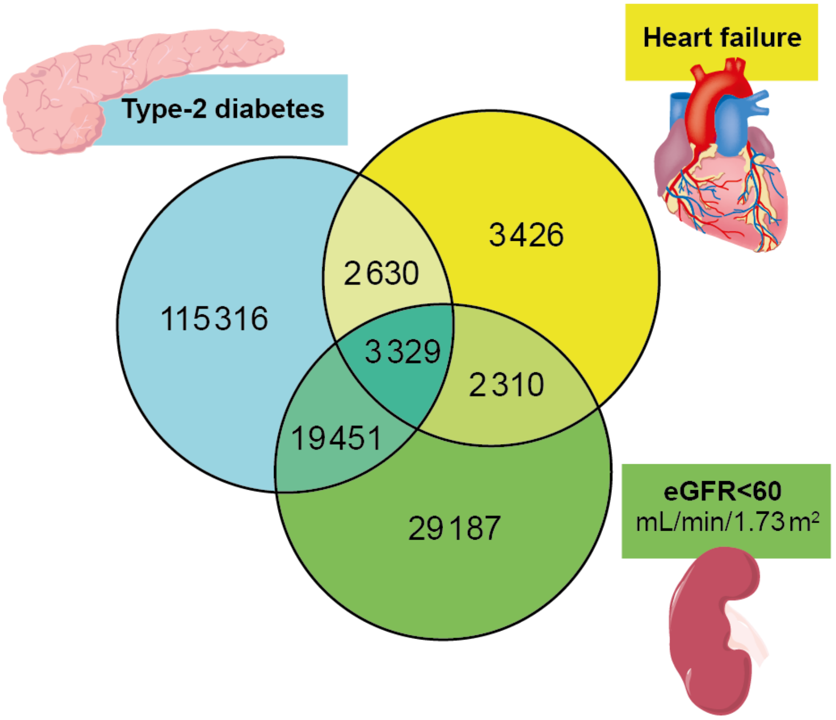
10.3.3.4 Find and use expected frequencies from Venn diagrams, two-way tables and tree diagrams

Common difficulties and misconceptions
When students are introduced to tree diagrams for the first time at Key Stage 4, they often recognise the need to multiply along the branches, without having a deep understanding of why this is necessary and when it is appropriate to do so. Provide opportunities for students to represent situations using tree diagrams, where the information needed for the branch labels is not explicitly given, to support them in understanding how a probability tree diagram models a situation. This is especially important given that, by this stage, students are expected to work with both independent and dependent events. Students need to identify whether the outcome of a previous event affects the probability of a subsequent event. A mixture of both types of event are used in the examples. Presenting frequency and probability tree diagrams simultaneously and asking students to make connections between these two representations, can help to expose mathematical structure and support understanding of how the probabilities in a probability tree diagram relate to the actual or expected frequencies. Students can often use probability trees to calculate probabilities, with little understanding of how the probabilities relate to a given situation. It is important that these representations are used to deepen students' ability to make these connections. This understanding is essential for students to be able to access questions that involve conditional probability, introduced for the first time in this key idea. Relating back to students' early work on fractions, using language such as 'whole' and 'part' to identify the relevant values for the denominator and numerator respectively, may support them in getting to grips with this challenging concept.

Students need to	**Guidance, discussion points and prompts**
Understand the relationship between probability and expected frequency *Example 1:* *Twins can be identical or non-identical.* *Identical twins can either share a placenta or have separate placentas; non-identical twins always have separate placentas.* *Roughly $\frac{2}{3}$ of twin pregnancies are non-identical.* *Approximately $\frac{3}{10}$ of identical twin pregnancies have separate placentas.* *a) Use this information to complete a probability tree, with the first set of branches identifying whether twins are identical, and the second identifying whether they share a placenta.* *Dr Akhtar is a twin specialist. So far in her career, she has delivered 385 sets of twins.* *b) Complete the Venn diagram below with the expected frequencies of each type of twin pregnancy that Dr Akhtar will have delivered.*	Medicine is a field of work that relies heavily on probability, so many of the examples from this key idea use medical statistics as their basis. The use of genuine contexts such as in *Example 1* is important for **deepening** students' awareness that the mathematics they are learning has real-world implications. This is particularly salient in part b, where students will have to consider how to round the expected values to give viable answers. Teachers need to be aware of the different **language** that is used to describe frequency and expectation in probability contexts. Teaching and examination materials might, for example, use 'experimental probability' and 'relative frequency' interchangeably. Students may find the latter term particularly challenging given that it uses the word 'frequency' but refers to a probability. Ensure that students understand that the word 'relative' implies that they are comparing a part (the frequency of a particular event) to a larger whole (the total number of trials). It is important to recognise the similarities and differences in the structure of each **representation**, so that students are aware of which parts of the Venn diagram relate to which parts of the probability tree. This is perhaps simpler in an example such as this, where there are only two branches in each tree, each corresponding to a 'yes' or 'no' answer to a question. Draw attention to how inclusion within a set in the Venn diagram is the same as responding

 *Dr Ibrahim works in another hospital. So far in her career, she has delivered 82 sets of twins, and 23 of these were identical twins sharing a placenta.* *c) Is this more or fewer pairs of identical twins sharing a placenta than you would expect? Use one of your representations to explain your answer.*	'yes' and therefore selecting the branch labelled 'Y' on the tree diagram.
Example 2: *The Venn diagram below this question comes from a medical research paper[3]. It shows the findings from a sample of 1,389,604 patients.* *It was found that 175,649 of the patients had one or more of three different conditions: type-2 diabetes, heart failure and chronic kidney disease.* *Dr Darby is a GP, treating all types of patients.* *a) What calculation would you need to do to find out the probability that one of her patients:* *(i) has diabetes and kidney disease* *(ii) has one or more of the three conditions* *(iii) has exactly one of the three conditions?* *Dr Oomatia treats patients with chronic kidney disease.* *b) What calculation would you need to do to find out the probability that one of his patients:* *(i) also has diabetes* *(ii) also has heart disease, but not diabetes* *(iii) only has chronic kidney disease?* *Dr Banting treats patients with diabetes.*	*Example 2* looks at how calculations can be constructed to find probabilities from frequencies. **Variation** is used to explore how the values needed change depending upon which population is being drawn from each time. This is particularly explicit in part d, which draws together all of the different instances of part (i) to compare what is the same and what is different. By keeping the conditions constant (kidney disease co-existing with diabetes), but varying the population each time, there is an opportunity to think deeply about how the values for the numerator and denominator are affected. Further exploration of conditional probability can be found in *Examples 7* and *8* below. Before each part-question, a sentence introducing a fictional medical professional is used to provide context for why each probability might be needed. This means that the **language** of the question is perhaps more 'wordy' than might be found in some textbook or examination questions on a similar topic. This is intentional: it is important that students have a sense of the 'why' when working with real data. This can give them a 'way in' to complex calculations, rather than just trying to unpick meaningless numbers. Students will be familiar with a Venn diagram **representation**, but they may not be aware of how actual researchers can use one to exemplify their findings. Here, the diagram has been simplified so that all of the sets are presented as the same size, but the original research paper uses three different size sets to represent the size of the populations. A link to the research paper can be found at the end of this document; students may be interested to see the Venn diagram in situ. In *Example 1*, the data used to summarise the situation was simplified for ease of calculation, whereas the data from *Example 2* is lifted directly from a research paper. It is important that students understand that data and probabilities can look 'messy' in

c) What calculation would you need to do to find out the probability that one of his patients: *(i) also has chronic kidney disease* *(ii) also has heart disease* *(iii) only has diabetes?* *All of your part (i) answers show the probability of a patient having both diabetes and chronic kidney disease.* *d) Why are each of your answers different?*	the real world, but equally important that this does not present a barrier to accessing new learning. Discuss with your team how you can achieve a balance between using real data and data that lead to straightforward calculation.

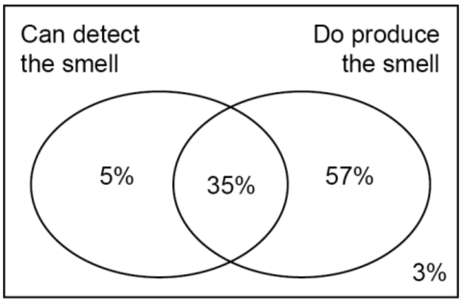


Use different representations to work with expected frequencies

Example 3:

Some people produce a particular smell in their urine after eating asparagus. Some people can smell this change, others cannot. There is significant overlap between the group that produces the smell and the group that detects it.

The following Venn diagram is based on researchers' estimates for the probability having one, both or neither of the traits[4]*.*



a) *How many people would you expect to both detect and produce the smell:*
   - (i) *in your class*
   - (ii) *in your year group*
   - (iii) *in your school?*

Marta lives with her mum, dad and two siblings. She says, 'I'd expect two members of my family to be able to detect the smell.'

b) *How has Marta worked out this expected frequency?*

Researchers have discovered that there is a gene responsible for the ability to detect the smell.

c) *What implications might this have for Marta's calculations?*

In the following series of examples, students are asked to work with different **representations** to consolidate their understanding of how probabilities can be used to find expected frequencies and vice versa. Noticing whether a particular representation causes difficulty can help inform future teaching. Teachers are encouraged to be curious about whether students are struggling with the wider concept of expected frequencies, or the reading and interpretation of a particular diagram. Working with more than one of these examples may help.

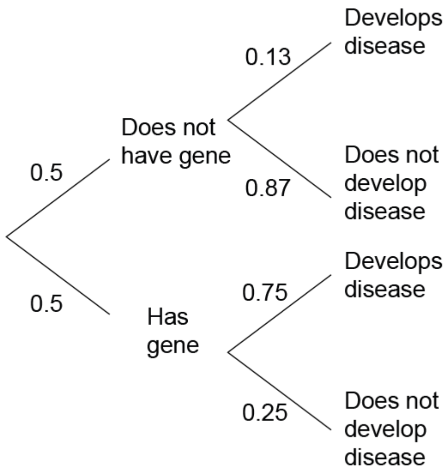
Part c is important for **deepening** students' understanding that probability calculations can provide a model, but that real life is often more complex, and so any models are likely to be a simplification of the situation. Here, the genetic component means that the probabilities for families that carry the gene will be different to the probabilities for the general population. Students' awareness of genetics will be informed by their own personal experiences, their awareness of current affairs, and their learning in science. Speak to colleagues in your science department to compare the science and maths curriculums, and familiarise yourself with students' prior learning on genetics before using a context such as this.

Real-life data do not have to be serious, and numerous examples can be found online of genuine research papers on more light-hearted topics. Using data that stimulate interest can be beneficial in terms of engagement and lesson memorability, but teachers also need to use their judgment about when this might distract from the intended learning point. This decision may be different for different classes in the same school. Consider taking time with your department to find interesting and engaging data and discuss whether it is appropriate for your current Key Stage 4 classes.

Example 4:

The lifetime chance of getting a disease is 13%. If a certain gene is inherited, this increases to 75%.

This situation is represented in the probability tree below.



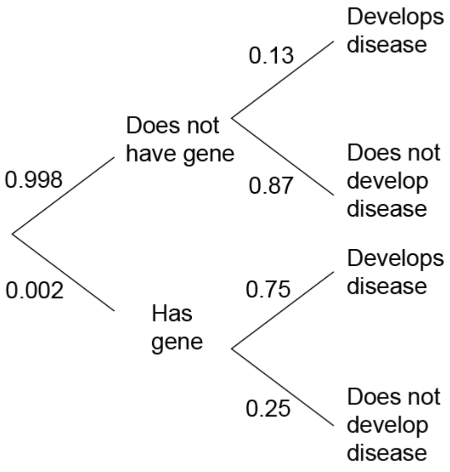
Emilia finds out that her mother has inherited the gene from her grandmother. In total, there are 21 women descended directly from her grandmother.

a) *How many of these women would you expect to inherit the gene?*

Emilia is one of four siblings. Two of her siblings have inherited the gene. Emilia says, 'That means I do not have it.'

b) *Is Emilia correct? Why or why not?*

In the general population, the chance of having the gene is 0.2%.



c) *If there are approximately 34.5 million women in the UK, how many would you expect to:*

Example 4 revisits the context from *Example 15* of 10.3.2.2, but here considers how the probabilities could be expected to affect a population rather than one individual. There is opportunity to explore misconceptions around expected outcomes of independent events, although this **language** is not explicitly used. For example, Emilia assumes that her siblings' inheritance of the gene affects her own. In fact, each subsequent birth is an independent event, each with the same likelihood of inheritance.

Parts a and d give an opportunity to talk about how to handle expected frequencies that are not integers, particularly in contexts where decimals are not possible, such as numbers of people. This can help with **deepening** students' understanding of the relationship between probability and expectation, and how the latter is not a guarantee but an estimation. Connections could therefore be made with estimates, and limits of accuracy, as the answer can be expressed as a range rather than a single value.

Part c offers an opportunity to model how students might interact with the **representation** when calculating probabilities, and to assess their depth of understanding. For example, highlighting along the branches can help students to identify the relevant values and calculations for different probabilities. When there are two different events modelled, as there are here ('Has the gene' and 'Develops disease') it can be tempting to see it as necessary to calculate along the branches for every possible probability. In fact, the probability for sub-part i just needs to be read from the first part of the probability tree, without any calculation. Indeed, it is arguably easier to calculate the number of women who have the gene $(0.002 \times 34\,500\,000)$ and then subtract this from the total number of women, rather than try to calculate $0.998 \times 34\,500\,000$. Students' willingness to spot efficiencies such as these is a strong indicator of their level of fluency when working with this diagram.

(i) have the gene

(ii) have the gene but not develop the illness in their lifetimes

(iii) develop the illness in their lifetime?

Dr Jones estimates that two or three of the patients registered at his surgery will have inherited the gene.

d) How many patients might be registered at his surgery?

Example 5:

Nessa is at her primary school fete. She has 20 p left to spend on a game and wants to have the best chance of winning a prize.

She watches the hook-a-duck and tombola for five minutes, and notices that each game has six winners in that time.

She says, 'There is an equal chance of winning both games, so it doesn't matter which I pick.'

a) Why might Nessa not be correct?

In fact, 42 people played hook-a-duck and 24 people played the tombola in that time.

b) Which game would you advise Nessa to choose? Why?

The tombola costs 20 p per go and the hook-a-duck costs 10 p per go.

c) Does this change your answer to part b? Why or why not?

Example 5 draws attention to what information is needed for likelihood to be ascertained. It is an opportunity to emphasise how the specific **language** reflects the structure of the mathematics: we know the frequency of winners, but without knowledge of how this relates to a total number of trials, we cannot determine the relative frequency.

Without a given **representation**, students are free to represent the situation in any way they find helpful. A tree diagram might be particularly helpful for part c, where it is revealed that Nessa can afford two goes of hook-a-duck, and so needs to compare the chance of winning over two trials, rather than just one.

Example 6:

A catering company is catering for an event and has carried out a survey of dietary requirements for the guests.

	Eat meat	**Don't eat meat**	**Totals**
Eat dairy	180	32	212
Don't eat dairy	35	3	38
Totals	215	35	250

Example 6 explores the relationship between probability and frequency using a two-way table **representation**. As with previous examples, students need to ascertain which of the five different totals are relevant here.

Students could feasibly work two ways: by finding 12% of the total of 250, to calculate an expected number of guests who do not eat meat, or by working out the relative frequency of guests who do not eat meat and comparing it to 12%. It may help with **deepening** students' understanding to explore both approaches and discuss how and why they can both be considered valid.

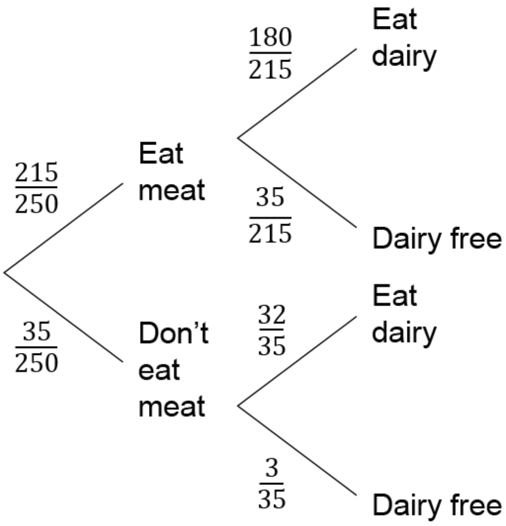
Part a asks if the data provided are representative of the wider population – spend time exploring students' understanding of this **language**. Do they make connections with the term 'representation'? It can be helpful to think about representations as a way of showing how one thing displays the same features as another. In

In the UK, 12% of people don't eat meat. a) *Are the guests at this event representative of this? How do you know?* *It is difficult to determine the percentage of people who do not eat dairy in the UK.* b) *Based on this data, what would be your estimate?*	the case of this example, students are asked whether the guests (a sample) display the same dietary features as the whole country (the population). The same set of data is used in several examples throughout this document, each time with a different pedagogical focus. Spend some department time looking at all of the related examples together: how does the emphasis change each time? Are any examples more, or less, effective than the others? Look at a data set together and generate as many different questions as you can, covering different key ideas from this theme. Are there any connections that could be made with concepts beyond this theme?
Work with conditional probability *Example 7:* *Kerry shuffles an ordinary deck of playing cards and selects a card at random.* a) *What is the probability of Kerry selecting:* (i) *a diamond* (ii) *a 4* (iii) *the 4 of diamonds* (iv) *a spade?* b) *If the first card Kerry selects is red, what is the probability that it is also:* (i) *diamond* (ii) *a 4* (iii) *the 4 of diamonds* (iv) *a spade?* c) *If the first card Kerry selects is a diamond, what is the probability that it is also:* (i) *red* (ii) *a 4* (iii) *the 4 of diamonds* (iv) *a spade?* d) *Write another set of questions with the same probabilities as parts b and c of this task.*	The **variation** in *Example 7* is such that parts a and b ask for the same probabilities to allow students to compare directly between them. The difference is that, in part b, we already know one condition: that the card is red. The simplicity of the questions here, keeping every other element the same, means that teachers can draw students' attention to what information they need to determine the probability each time. Students should be aware that, when expressing the probability as a fraction, the denominator will change according to the different conditions in parts a to c, and the numerator will change according to the different characteristics in sub-parts i to iv. Consider the **language** commonly used for conditional probability questions. It might be helpful to use a deliberate emphasis on words such as 'when' and 'if' to support students to understand which value is being used as the 'whole' when calculating the probability. For example, the question, 'If the first card is red, what is the probability that it is also a diamond?' could be rephrased as, 'What is the probability of diamonds when the first card selected is red?' Students could mistakenly assume that the denominator is always the frequency referred to in the first part of the sentence, and so exemplifying different sentence structures could help students to avoid this misconception. Part d offers an opportunity for **deepening** students' understanding by asking them to work backwards. The guidance for *Example 3* from 10.3.3.2 explores this pedagogical structure further. The question prompt, '*What is the same and what is different?*' can be useful for working with carefully varied examples such as this one. With your team, pick pairs of sub-parts from this example to compare using this question structure. Which pairs are most useful to compare? Why?

Example 8:

A catering company is catering for an event and has carried out a survey of dietary requirements for the guests.

Olly uses a tree diagram to find the probability that a guest who doesn't eat meat is also dairy free.



Olly multiplies along the branches and tells Petra that the probability is:

$$\frac{35}{250} \times \frac{3}{35} = \frac{3}{250}.$$

Petra disagrees and says that the probability can be read off as $\frac{3}{35}$.

a) *Who is correct? Explain how you know.*

The company changes the menu so that the dairy-free option is also vegetarian.

c) *Create a new probability tree where the guests are first divided by whether they eat dairy.*

d) *Which tree diagram would be most useful for working out the probability of:*

  (i) *a guest being dairy free?*

  (ii) *a guest not eating meat?*

  (iii) *a guest who eats dairy not eating meat?*

  (iv) *a guest who eats meat not eating dairy?*

In *Example 8*, an emphasis is placed on identifying whether the whole population, or a subset of the population, is being considered. Explore the ways in which probability tree diagram **representations** can be used to reveal mathematical structure. Creating, and then working with, two differently-structured probability trees for the same data should help students to build a deep and connected understanding of this representation, rather than just memorising processes such as 'multiply along the branches'.

Distinguishing between the need to consider all 250 guests and just a subset (in this case, those guests who don't eat meat) can be quite subtle. An example such as this is important in **deepening** students' understanding. Representing the survey results in a two-way table (as shown with the same set of data in *Example 6*) may help some students to recognise the difference between multiplying along the branches (when we are interested in one of the 250 guests selected at random) and referring to the fraction on a specific branch (when we are only interested in the 35 guests who don't eat meat).

Using multiple representations together can help to deepen students' understanding of a particular representation. Are there any other representations that might be helpful to consider here? In what ways might a Venn diagram, for example, provide additional insight into the structure of a probability tree diagram?

Collaborative planning

Although they may provoke thought if read and worked on individually, the materials are best worked on with others as part of a **collaborative professional development** activity based around planning lessons and sequences of lessons.

If being used in this way, is important to stress that they are not intended as a lesson-by-lesson scheme of work. In particular, there is no suggestion that each key idea represents a lesson. Rather, the fine-grained distinctions offered in the key ideas are intended to help you think about the learning journey, irrespective of the number of lessons taught. Not all key ideas are of equal weight. The amount of classroom time required for them to be mastered will vary. Each step is a noteworthy contribution to the statement of knowledge, skills and understanding with which it is associated.

Some of the key ideas have been extensively exemplified in the guidance documents. These exemplifications are provided so that you can use them directly in your own teaching but also so that you can critique, modify and add to them as part of any collaborative planning that you do as a department. The exemplification is intended to be a starting point to catalyse further thought rather than a finished ‘product’.

A number of different scenarios are possible when using the materials. You could:

- Consider a collection of key ideas within a core concept and how the teaching of these translates into lessons. Discuss what range of examples you will want to include within each lesson to ensure that enough attention is paid to each step, but also that the connections between them and the overall concepts binding them are not lost.
- Choose a topic you are going to teach and discuss with colleagues the suggested examples and guidance. Then plan a lesson or sequence of lessons together.
- Look at a section of your scheme of work that you wish to develop and use the materials to help you to re-draft it.
- Try some of the examples together in a departmental meeting. Discuss the guidance and use the PD prompts where they are given to support your own professional development.
- Take a key idea that is not exemplified and plan your own examples and guidance using the template available at Resources for teachers using the mastery materials | NCETM.

Remember, the intention of these PD materials is to provoke thought and raise questions rather than to offer a set of instructions.

Solutions

Solutions for all the examples from *Theme 10 Statistics and probability* can be found here:
https://www.ncetm.org.uk/media/r2idmejd/ncetm_ks4_cc_10_solutions.pdf

Data sources

1 Our World in Data. (n.d.). *Affiliation of research teams building notable AI systems, by year of publication*. [online] Available at: https://ourworldindata.org/grapher/affiliation-researchers-building-artificial-intelligence-systems-all.

2 Boniel-Nissim, M., Marino, C., Galeotti, T., Blinka, L., Ozoliņa, K., Craig, W., Lahti, H., Wong, S., Brown, J., Wilson, M., Inchley, J. and Van Den Eijnden, R. (2024). *A focus on adolescent social media use and gaming in Europe, central Asia and Canada*. [online] Available at: https://iris.who.int/items/b7538d5f-d9ba-455f-ae11-80b28a43c76a

3 Schechter, M., Cheli Melzer Cohen, Ilan Yanuv, Rozenberg, A., Chodick, G., Johan Bodegård, Leiter, L.A., Verma, S., Lambers, J., Karasik, A. and Ofri Mosenzon (2022). Epidemiology of the diabetes-cardio-renal spectrum: a cross-sectional report of 1.4 million adults. *Cardiovascular Diabetology*, 21(1). Doi: https://doi.org/10.1186/s12933-022-01521-9.

4 Agata Blaszczak-Boxe (2016). Asparagus Pee? Why Only Some People Smell It. [online] Live Science. Available at: https://www.livescience.com/57199-why-some-people-cant-smell-asparagus-pee.html [Accessed 30 Jul. 2025].

Markt, S.C. and Mucci, L.A. (2017). Authors' reply to Rishniw. *BMJ*, 2910: https://doi.org/10.1136/bmj.j2910.