

Examiners' report

A LEVEL

PSYCHOLOGY

OCR Level 3 Advanced GCE in Psychology

H567

For first teaching in 2015

H567/01 Summer 2025 series

Contents

Introduction	3
Paper 1 series overview	4
Section A overview	5
Question 3	5
Question 14	6
Question 16	6
Question 19 (a)	7
Question 19 (b)	7
Section B overview	8
Question 20	8
Question 21*	9
Question 22	12
Question 23	12
Question 24	13
Question 25	14
Section C overview	15
Question 26 (a)	15
Question 26 (b)	17
Question 27	18
Question 28 (a)	18
Question 28 (b)	19
Question 28 (c)	20
Question 29 (a)	20
Question 29 (b)	21
Question 30	22

Introduction

Our examiners' reports are produced to offer constructive feedback on candidates' performance in the examinations. They provide useful guidance for future candidates.

The reports will include a general commentary on candidates' performance, identify technical aspects examined in the questions and highlight good performance and where performance could be improved. A selection of candidate responses is also provided. The reports will also explain aspects which caused difficulty and why the difficulties arose, whether through a lack of knowledge, poor examination technique, or any other identifiable and explainable reason.

Where overall performance on a question/question part was considered good, with no particular areas to highlight, these questions have not been included in the report.

A full copy of the question paper and the mark scheme can be downloaded from [Teach Cambridge](#).

Would you prefer a Word version?

Did you know that you can save this PDF as a Word file using Acrobat Professional?

Simply click on **File > Export to** and select **Microsoft Word**

(If you have opened this PDF in your browser you will need to save it first. Simply right click anywhere on the page and select **Save as . . .** to save the PDF. Then open the PDF in Acrobat Professional.)

If you do not have access to Acrobat Professional there are a number of **free** applications available that will also convert PDF to Word (search for PDF to Word converter).

Paper 1 series overview

This is the first of the three examined components for the OCR A Level Psychology qualification. There was a wide range of responses, suggesting that the paper differentiated fairly. More successful candidates wrote more extended and detailed responses that were contextualised and made good use of appropriate terminology. Some candidates found it difficult to use terminology appropriately and used concepts such as validity and reliability with some confusion. Those candidates also struggled with contextualising their responses – they simply rewrote the question stem/scenario given to introduce their answers throughout sections B and C.

As in the previous series, it was necessary for candidates to have some practice in the design and implementation of their own practical activities. This should have reinforced their knowledge and understanding of research methods in general, as well as some of the specific terms and concepts they could be assessed on. This would help candidates to comment on how conducting their own research has helped in the planning of novel research presented in this examination.

It is important to note that research methods should be reinforced through the core studies, since section A might require candidates to identify the research methods used in any of the core studies. Finally, candidates should be encouraged to use examples to illustrate key points that they make. This would help them to convey their understanding better.

Contextualising responses

Rewriting the question stem/scenario detailed to introduce answers throughout sections B and C will not get context marks unless evaluation points are applied to the scenario directly. The mere presence of 'context words' in the response is not sufficient to contextualise response.

Candidates who did well on this paper generally:	Candidates who did less well on this paper generally:
- contextualised their responses - had a good understanding of the correlational method, sampling techniques, levels of data and ethics and were able to provide strengths and/or weaknesses for these in their responses - wrote extended responses appropriate to the number of marks each question was worth - understood the difference between findings and conclusions - had a good understanding of how to calculate Chi-Square and how to interpret a significance statement - had a good understanding of factors that can affect validity and reliability.	- showed inconsistency in contextualising responses - had some understanding of the correlational method, sampling techniques, levels of data and ethics although this was inconsistent - wrote less extended responses sometimes not appropriate to the number of marks each question was worth - frequently lacked understanding of how to use Chi-Square, e.g. adding extra steps in the calculations - had some understanding of factors that can affect validity and reliability, but some confusion between the two was present.

Section A overview

There was a good knowledge and understanding shown of the research methods, descriptive and inferential statistics as well as mathematical calculations. In this section, candidates need to be prepared to identify research methods concepts in the core studies.

Overall feedback on multiple-choice questions

Candidates answered most of the multiple-choice questions well. Incorrect responses were varied. The notable exceptions are outlined below.

Question 3

- 3 A study investigated if there was a difference in memory for a list of 30 words. The same participants were assessed at 9am and again at 9pm. Memory was measured as a score out of 30.

Which inferential statistical test would be used to analyse the data from this study?

- A Binomial Sign test
- B Chi-square
- C Mann-Whitney U test
- D Wilcoxon Signed Ranks test

Your answer

[1]

Answers to this question varied, with many candidates choosing Mann-Whitney, which was incorrect.

Question 14

14 Which of these is the correct way to reference Milgram's (1963) study of obedience using the Harvard system?

- A** Milgram, S. (1963) Behavioral study of obedience *Journal of Abnormal and Social Psychology*, 67, (4), 371–8
- B** Milgram, S. Behavioral study of obedience (1963) *Journal of Abnormal and Social Psychology*, 67, (4), 371–8
- C** S. Milgram, (1963) Behavioral study of obedience *Journal of Abnormal and Social Psychology*, 67, (4), 371–8
- D** S. Milgram, Behavioral study of obedience (1963) *Journal of Abnormal and Social Psychology*, 67, (4), 371–8

Your answer

[1]

Many candidates answered this question incorrectly – incorrect answers were varied.

Question 16

16 When report writing, how are the details of previous researchers' work recorded in the references section?

- A** alphabetically by researcher's surname
- B** date order (with the most recent publication first)
- C** most frequently cited research presented first
- D** relevance to the research being reported (with most relevant first)

Your answer

[1]

Many candidates answered this question incorrectly – incorrect answers were varied.

Question 19 (a)

19 From Bandura et al.'s (1961) study into transmission of aggression:

(a) Which variables did the researchers correlate to check inter-rater reliability?

- A** aggression scores of girls and boys
- B** number of aggressive acts by models and children
- C** number of observers and number of aggressive acts observed
- D** ratings of aggression by two observers

Your answer

[1]

Question 19 (b)

(b) How was the setting controlled in the final observation room?

- A** the children observed the same model
- B** the children were only allowed to play for 2 minutes
- C** the observers stood in the same place in the room
- D** the toys were placed in the same position for each participant

Your answer

[1]

Questions 19 (a) and (b) were incorrectly answered by many candidates. It is important to make sure that candidates are prepared to identify research methods concepts in the core studies.

Section B overview

Overall, there was a reasonable understanding shown of the correlational research, including sampling techniques, strengths and weaknesses and controls. Candidates were able to identify factors affecting validity and outline appropriate ethical considerations that should be accounted for in the study presented.

To improve, candidates should practice operationalising variables for correlations as these proved particularly challenging. Candidates should also focus on detailing how the required features in Question 21 would be enacted/carried out. It is also important to contextualise responses to the scenario referred to, e.g. conversations differ to public speech therefore candidates should strive to give more specific context.

Question 20

Um, erm and ah

Disfluencies are disruptions in the flow of spoken language, including stuttering and hesitations, such as 'um', 'erm' and 'ah'. These often indicate emotions, such as feelings of anxiety and distress. A psychologist wants to investigate this further by conducting a correlation study to see if there is a relationship between the disfluencies a person makes and how anxious they feel while making a public speech.

20 Write a null hypothesis for this study.

.....

.....

.....

.....

.....

..... [3]

Overall, this question was answered well – candidates showed more confidence in writing suitable hypotheses.

Although many candidates got at least 1 mark for 'no significant correlation', there were still many candidates who did not operationalise variables. Those that did not get any marks wrote an experimental (i.e. referred to a difference) or a directional hypothesis.

Question 21*

21* Explain how you would conduct a correlation study to investigate if there is a relationship between the disfluencies a person makes and how anxious they feel while making a public speech. Justify your decisions as part of your explanation.

You must refer to:

- how you would use self-selected sampling to obtain participants for the study
- how you would operationalise the variable 'disfluencies'
- how you would operationalise the variable 'anxiety'
- the control of **one** extraneous variable.

You should use your own experience of practical activities to inform your response.

[15]

The standard of responses was reasonable this series, with most candidates being able to identify and describe each of the required features (RF). However, most candidates found it difficult to get the higher band marks. The most successful responses were characterised by writing a separate paragraph relating to each of the RFs. First, the responses demonstrated understanding of what was involved and how to address it for the research presented. Next, they justified the decisions made regarding how to address it. Finally, the responses drew on the candidates' own experiences of conducting research. All this needed to be discussed in context to get marks. There was much variation in how candidates demonstrated knowledge and understanding of each of the individual RFs. Often, candidates did not provide enough detail.

RF1 (the use of self-selected sampling) was generally addressed well. However, some candidates did not mention how they obtained the final sample. Successful responses were characterised by describing how the study will be advertised (e.g. poster), how participants will contact the researcher (e.g. researcher's email on the poster) and how the researcher will end up with a final sample (e.g. first 20 to email will take part in the study). Some candidates did not justify their choices in the relevant context as evaluations were general (e.g. a strength is that it is easy to obtain a sample) rather than unique to self-selected sampling (e.g. less likely to withdraw from the study as they have put themselves forward).

RF2 (operationalising the variable 'disfluencies') seemed to be the most challenging, with many candidates forgetting that at least an ordinal level of data is required for a correlation. This meant that many responses used nominal data as they referred to the use of event sampling, tallying in behavioural categories and multiple observers gathering data. Those could be made relevant if candidates would calculate the average of disfluencies or total number of disfluencies (e.g. find an average number of disfluencies using data from all observers), however many of them did not do so.

RF3 (operationalising the variable 'anxiety') was answered well with specific examples such as self-reported levels of anxiety during a public speech on a scale of 1-10. Those who explained the ends of the scale and referred to public speech were most likely to be placed in the good (L4) level for that RF. However, candidates did not get marks if they indicated more than one measure or if they have misnamed the scale (e.g. a Likert scale on a scale of 1-10). Justifications for this were varied and some were not in context (e.g. quantitative data is easy to compare) or not expanded on (e.g. easy to analyse).

RF4 (a control of one extraneous variable) had the most varied responses, with some candidates having trouble choosing and explaining appropriate controls. More successful responses made references to controlling the situational variables in terms of noise, time of the day or size of the audience. Some candidates chose to control personal characteristics, e.g. language barriers, confidence levels, previous experience of public speaking, but this was not creditworthy in this instance, as a range of participants are needed for a correlation to fully explore levels of anxiety and disfluencies. Justifications were often linked to increased validity and/or reliability with more successful responses explaining why the control would lead to increased reliability and/or validity. Less successful responses confused the two concepts or simply stated the link to validity/reliability with no further elaboration.

Most candidates made reference to their own research. However, for some of them, this was not explicit (it was impossible to determine what their study was about or how their chosen RF was used) and did not offer any link to their justification.

Exemplar 1

In my study I would operationalise the variable of 'disfluencies' by counting the number of times a person says either 'um', 'erm' or 'ah' throughout ~~the~~ 30 minutes ^{of a} public speech. This will provide me with ~~more reliable~~ ^{consistent} data as it can be standardised for every participant. By recording the number of disfluencies in a 30 minute period, the data collected will be more reliable as it was consistently recorded in the same way for everyone, rather than using the whole length of the speech, as this may vary for each participant, making the number of disfluencies ~~less reliable~~ less reliable. In my practical activity investigating the correlation between tiredness and reaction time, I used the same ruler for all participants so reaction time ~~can~~ could be measured using the ruler drop test in a standardised way to increase reliability. Hence why I will operationalise the ~~variable~~ variable of 'disfluencies' in this study in a similarly standardised way.

This exemplar shows a Level 4 (good) example of operationalising 'disfluencies' correctly – the candidate clearly operationalised the co-variable (including time frame of measurement and reference to public speech) and justified their response, all of which was done in context.

Assessment for learning



The most successful responses are characterised by addressing each RF separately. Within each paragraph, candidates should:

- demonstrate an understanding of RF and how to address it in context of the scenario given
- justify the decisions made in context of the scenario given
- draw on their own experiences of conducting research and demonstrate how it informed their planning of the currently proposed research.

Question 22

22 Outline **one** strength and **one** weakness of conducting this study using the correlation technique.

.....

.....

.....

.....

.....

..... [6]

Candidates who got full marks were able to identify strengths and weaknesses in context and explained why this would be considered a strength or a weakness. Less successful responses either did not include context or did not comment on why something is a strength or a weakness. Some candidates would simply state 'therefore more/less valid and/or reliable' rather than being specific. Those who got 0 marks focused on generic evaluation points that could apply to any research method (e.g. cheap/easy).

Assessment for learning



Questions requiring candidates to outline one strength and/or weakness tend to be answered well if they follow point, explain, context structure.

Question 23

23 Outline **one** weakness of using self-selected sampling in this study.

.....

.....

.....

.....

.....

..... [3]

Candidates were generally successful with this question, with most stating that those who volunteer might be less anxious or that volunteers tend to share similar personality characteristics. Less successful responses did not apply context effectively.

Question 24

24 Outline **one** ethical consideration that could be taken into account in this study.

.....

.....

.....

.....

.....

..... [2]

Candidates generally answered this question well – more successful responses showed good knowledge in context. Some candidates mixed up issues, for example referencing respect but then outlining protection from harm.

Question 25

25 Outline **two** ways the validity of the data collected in this study could have been affected.

1

.....

.....

.....

.....

.....

2

.....

.....

.....

.....

.....

[6]

More successful responses identified possible reasons for validity being affected in context as well as elaborated on them to explain why those features would help with increasing/decreasing validity. There were many candidates that got fewer than 6 marks for mixed reasons. Firstly, candidates sometimes missed out context. Secondly, candidates were unable to explain how this affected the validity, i.e. increased/decreased.

Some candidates found this question challenging and focused on population and ecological validity without any reference to data, which meant that they were unable to move past 'weak attempt at outline (whether in context or not)' which was worth 1 mark for each way.

Candidates who did not get any marks on this question usually referred to reliability instead of validity.

Section C overview

A good understanding of pie charts and inferential statistics was shown in this section. In order to improve, candidates should practice writing conclusions as well as differentiate between levels and types of data in terms of their unique strengths and weaknesses.

Question 26 (a)

Stress

Stress can affect everyone, but not always in the same way and it is not always caused by the same thing.

A psychologist conducted a study of 120 people (60 males, 60 females) aged 16 to 65 using the self-report method to investigate people's experience of stress.

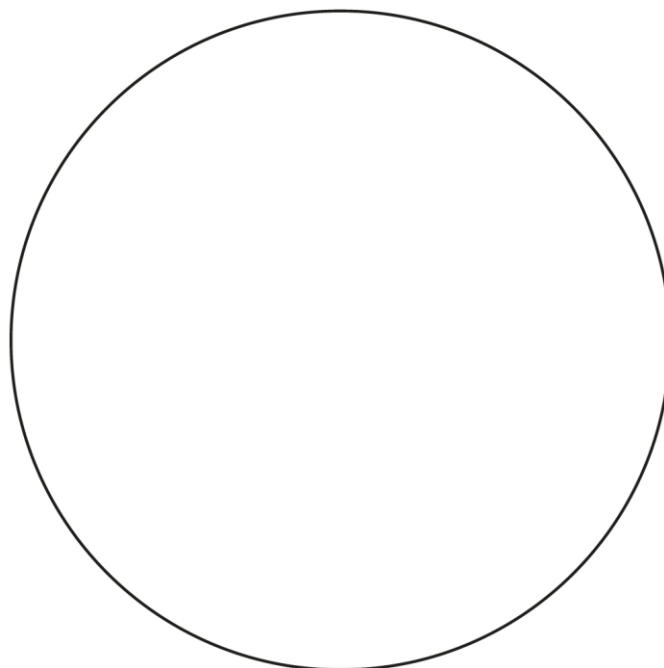
Some of the data collected is presented in **Table 1** and **Table 2**.

Table 1

Number of times each category was selected when asked 'Which one of the following do you find the most stressful?'				
health	finances	relationships	work	family
33	24	12	42	9

Table 2

Number of 'yes' and 'no' responses when asked 'Do you think stress can be a good thing?'		
	Yes	No
Males	32	28
Females	14	46

26**(a)** Sketch a fully labelled pie chart showing the data displayed in **Table 1**.**[4]**

Many candidates missed at least 1 mark. The most common mistakes included miscalculating (or not calculating at all) the proportions or missing the title. Many candidates left the calculation on the table and did not transfer them over to the pie chart – while this did not matter in this series, it is not best practice.

Question 27

- 27** Use the data in **Table 2** to calculate the percentage of people who answered 'yes' to the question 'Do you think stress can be a good thing?'
Show your working.
Write your answer to **two** significant figures.

Table 2

Number of 'yes' and 'no' responses when asked		
'Do you think stress can be a good thing?'		
	Yes	No
Males	32	28
Females	14	46

.....

.....

.....

.....

.....

..... [3]

This was generally answered well, but more emphasis has to be placed on reading the question carefully, as many candidates missed the requirement to present the answer to two significant figures.

Question 28 (a)

- 28**
(a) Outline **one** reason why the Chi-square test would be the appropriate inferential statistical test to analyse the data in **Table 2**.

.....

.....

.....

..... [2]

This question was generally answered well. However, many candidates did not contextualise their response, i.e. identified nominal data but then stated male/female. Candidates should be aware that the data is related to the DV rather than the IV in order to match the correct context to the feature.

Question 28 (b)

(b) **Table 3** shows the data from Table 2 with expected frequencies (E values) provided in bold and in brackets.

Table 3

Number of 'yes' and 'no' responses when asked		
'Do you think stress can be a good thing?'		
	Yes	No
Males	32 (23)	28 (37)
Females	14 (23)	46 (37)

Formula for Chi-Square test

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Use the formula provided to calculate the value of Chi-square for this data.
Use **two** decimal places for all calculations.
Show your working.

.....

.....

.....

.....

.....

.....

..... **[5]**

This was generally answered well. Errors typically came from rounding to two decimal places incorrectly or adding an extra step (square rooting 11.42).

Question 28 (c)

- (c) Explain what $p < 0.05$ would mean as part of the significance statement for the analysis of this data.

.....

.....

.....

.....

.....

..... [3]

This question generally was not well-answered, as there was a lot of hypothetical language used (e.g. 'if the critical value is higher...then it could be...').

Candidates often missed a mark by not referring to the results as showing 'less than 5 percent probability of results being due to chance' or not writing this statement clearly enough.

Question 29 (a)

29

- (a) Outline **one** way that using a self-report method **increases** the reliability of the data collected in this study.

.....

.....

.....

.....

.....

..... [3]

Most candidates were able to get 1 or 2 marks on this question by identifying a relevant point (i.e. mostly standardisation) and putting it in context. Not many were able to get the final mark as they did not elaborate on consistency of data or ability to replicate the study to check for said consistency. Those who referenced sample size had trouble going beyond 1 mark as they were unable to explain how this increases reliability (therefore provided a weak outline, whether in context or not).

Many candidates identified points relevant to validity rather than reliability (e.g. use of primary data).

Misconception



Being able to replicate research does not automatically increase reliability. Replication only allows us to test for consistency; there is a chance that we replicate research and still find inconsistent (i.e. unreliable) results.

Question 29 (b)

- (b) Outline **one** way that using a self-report method **decreases** the reliability of the data collected in this study.

.....

.....

.....

.....

.....

..... [3]

This question proved more difficult in comparison to Question 29 (a). More successful responses referred to overlapping categories (relationships/family), which could lead to inconsistent answers over time or misinterpretation by participants. However, many responses did not differentiate between validity and reliability and referenced lying or social desirability, which is not creditworthy.

Exemplar 2

the level of stress someone feels is subjective
and could change from day to day, if the
same participants were asked the same question
on a different day their answer could be different
making it not very consistent

..... [3]

This response got 3 marks. The candidate identified one way that would decrease reliability and explained why this would lead to lower reliability (different answers over time = less consistency). All of this was written in the context of the study.

Question 30

30 Outline **one** strength and **one** weakness of using nominal data in this study.

.....

.....

.....

.....

.....

..... [6]

Many candidates gave advantages and disadvantages that were not specific to nominal data and instead applied more generally to quantitative data. This was given a maximum of 2 marks as both were classed as an attempt (whether in context or not). More successful responses referred to pie charts/headcount/frequency tallies that made strengths more specific to nominal data. Even more successful responses made references to other levels of data in their evaluation, e.g. since it is just a headcount it does not tell us the extent of people's feelings towards stress like ordinal data can.

Exemplar 3

On the other hand, nominal data may not be as good for this study about stress because it doesn't allow you to find ^{some} measures of central tendency or dispersion (eg variance) ^(eg mean) for the ~~stres~~ results of where stress comes from. This would mean averages and ~~most~~ a lot of analysiss can't be carried out to find out more about ~~wh~~ responses to what causes stress. [6]

This response got 3 marks (NB. as per mark scheme, one strength/weakness can get up to 3 marks). The candidate identified a weakness and explained why this would be a weakness. All of this was written in the context of the study.

Supporting you

Teach Cambridge

Make sure you visit our secure website [Teach Cambridge](#) to find the full range of resources and support for the subjects you teach. This includes secure materials such as set assignments and exemplars, online and on-demand training.

Don't have access? If your school or college teaches any OCR qualifications, please contact your exams officer. You can [forward them this link](#) to help get you started.

Reviews of marking

If any of your students' results are not as expected, you may wish to consider one of our post-results services. For full information about the options available visit the [OCR website](#).

Access to Scripts

We've made it easier for Exams Officers to download copies of your candidates' completed papers or 'scripts'. Your centre can use these scripts to decide whether to request a review of marking and to support teaching and learning.

Our free, on-demand service, Access to Scripts is available via our single sign-on service, My Cambridge. Step-by-step instructions are on our [website](#).

Keep up-to-date

We send a monthly bulletin to tell you about important updates. You can also sign up for your subject specific updates. If you haven't already, [sign up here](#).

OCR Professional Development

Attend one of our popular professional development courses to hear directly from a senior assessor or drop in to a Q&A session. Most of our courses are delivered live via an online platform, so you can attend from any location.

Please find details for all our courses for your subject on **Teach Cambridge**. You'll also find links to our online courses on NEA marking and support.

Signed up for ExamBuilder?

ExamBuilder is the exam paper builder platform for a range of our qualifications. [Find out more](#).

Free for all OCR centres with an Interchange account, it allows unlimited users per centre. We require an [Interchange](#) username to verify your centre's users for ExamBuilder.

If you do not have an Interchange account, please contact your centre administrator (usually the Exams Officer) to request a username.

Once you have an Interchange username, use the form on [this page](#) to sign up for ExamBuilder, or ask an existing user at your centre to create an ExamBuilder account for you.

Active Results

Review students' exam performance with our free online results analysis tool. It is available for all GCSEs, AS and A Levels and Cambridge Nationals (examined units only).

[Find out more](#).

Tell us what you think

Your feedback plays an important role in how we develop, market, support and resource qualifications now and into the future. We want you and your students to enjoy and get the best out of our qualifications and resources, but to do that we need your honest opinions to tell us whether we're on the right track or not.

You can email your thoughts to resources.feedback@ocr.org.uk or visit our [feedback page](#) to learn more about how you can help us improve our resources.



Designing and testing in [collaboration with you](#) and your students



Helping young people develop an [ethical view of the world](#)



Equality, diversity, inclusion and belonging (EDIB) are [part of everything we do](#)

Contact the team at:

☎ 01223 553998

🖥 ocr.org.uk

To stay up to date with all the relevant news about our qualifications, register for email updates at ocr.org.uk/updates

Visit our Online Support Centre at support.ocr.org.uk



OCR is part of Cambridge University Press & Assessment, a department of the University of Cambridge.

OCR is a Company Limited by Guarantee and an exempt charity. Registered in England. Registered office: The Triangle Building, Shaftesbury Road, Cambridge, CB2 8EA. Registered company number: 3484466.

We operate academic and vocational qualifications regulated by Ofqual, Qualifications Wales and CCEA as listed in their qualifications registers.

We are committed to providing a fully accessible experience across all our products, platforms, and websites. Find out more about our [accessibility standards](#).

© OCR 2025. All rights reserved. We retain the copyright on all our publications. However, our registered centres are permitted to copy and distribute our material for their own internal use, in line with any specific restrictions detailed in the publication. Find out more about our [copyright policies](#).

We update our publications regularly so please check you have the most up-to-date version.

We cannot be held responsible for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and do not guarantee that any content on such websites is, or will remain, accurate or appropriate.

When we update our specifications, you'll see a new version number and a summary of the changes. While we do our best to reflect these changes in all associated resources on [Teach Cambridge](#), if you notice any discrepancies, please refer to the latest specification on our website and [let us know](#).

Our resources do not represent any teaching method we expect you to use. We cannot be held responsible for any errors or omissions in our resources.