

Date: 18th August 2020

Our Reference: FOIA-2020-089

RE: Your request for information under the Freedom of Information Act 2000

I write in response to your Freedom of Information Act 2000 (or 'FoIA 2000') request dated **21/07/2020** in which you requested:

"The Day One Report".

Decision

After conducting careful searches for any information relevant to your request I can confirm that there is information held.

Please see the attached Day One Report which will be uploaded onto the College's website tomorrow

Your rights are provided at the bottom of this letter.

Yours sincerely,

Kate Kaufman | Legal Researcher
Information Management and Legal Team
College of Policing

Email: FOI@college.pnn.police.uk

Website: www.college.police.uk

Your right of review

Under the Freedom of Information Act 2000 you have a right to request an internal review if you are dissatisfied with our handling of your request. Review requests should be made in writing (by email or post) within **40 working days** from the date of our original response. We will aim to respond to your review request within **20 working days**.

The Information Commissioner's Office (ICO)

If, after lodging a review request you are still dissatisfied, you may raise the matter with the ICO. For further information you can visit their website at <https://ico.org.uk/for-the-public/official-information/>. Alternatively, you can contact them by phone or write to them at the following address:

Information Commissioner's Office

Wycliffe House

Water Lane

Wilmslow

Cheshire

SK9 5AF

Phone: 0303 123 1113



College of
Policing

college.police.uk

Day One Assessment Centre

Analysis of Pilot Assessment Centre Data

August 2020



© College of Policing Limited (2020)

This publication is licensed under the terms of the Non-Commercial College Licence v1.1 except where otherwise stated. To view this licence visit http://www.college.police.uk/Legal/Documents/Non_Commercial_College_Licence.pdf

Where we have identified any third-party copyright information, you will need to obtain permission from the copyright holders concerned.

Any enquiries regarding this publication or to request copies in accessible formats please contact us at contactus@college.pnn.police.uk

Contents

Tables	4
1. Introduction	5
2. Overall analysis	7
3. Exploration of biodata interactions	16
4. Exercise performance	19
5. Discussion	23
6. Next steps	28
7. Logistical regression analysis	29
Appendix A: Diversity by design	31
Appendix B: Overview of Functional Skills Level 2 English Language	32
Appendix C: Logistical regression methodology	33
Appendix D: Interactions within logistical regression model	34
Appendix E: Effect size labels	36

Tables

Table 1 – Overall assessment centre success rate	7
Table 2 – Overall assessment centre success rate by force	7
Table 3 – Attendance and success rate by sex	8
Table 4 – Attendance and success rate by first language	8
Table 5 – Attendance and success rate by disability	9
Table 6 – Attendance and success rate by age (banded)	9
Table 7 – Attendance and success rate by ethnicity (2+1) with AIRs	10
Table 8 – Attendance and success rate by ethnicity (5+1) with AIRs	11
Table 9 – Attendance and success rate by ethnicity with AIRs	12
Table 10 – Attendance and success rate by ethnicity with AIRs (English as first language only)	14
Table 11 – Ethnicity representation of overall candidate pool and successful candidates	16
Table 12 – Candidates’ ethnicity (2+1) by age	16
Table 13 – Candidates’ ethnicity (2+1) by sex	17
Table 14 – Candidates’ ethnicity (2+1) by first language	17
Table 15 – Overall assessment centre success rate by ethnicity (2+1) by force	18
Table 16 – Exercise analysis by ethnicity (low technology pilot forces only – Thames Valley Police, Surrey and Sussex)	20
Table 17 – Exercise analysis by ethnicity	21
Table 18 – Exercise analysis by sex	22

1 Introduction

The Day One Assessment Centre has been designed through collaboration between the College of Policing, the Metropolitan Police Service (MPS) and a third-party commercial supplier, to replace the SEARCH® National Recruit Assessment Centre. The Day One Assessment Centre has been piloted across four pilot forces between April 2018 and November 2019.

This report provides analysis of the candidates who attended the Day One Assessment Centre during the pilot period. A total of 18,150 candidates attended the assessment centre between 4 April 2018 and 30 November 2019.

The analysis set out in this report is primarily descriptive statistics, including evaluation of adverse impact ratios (AIRs) and effect sizes (an explanation of AIRs and effect sizes is given in the **Discussion** section of this report) for key biographical information, particularly focusing on candidates' ethnicity.

Insights into data analysis conducted to date are covered in the **Discussion** section and key milestones are set out in the **Next steps** section.

How was Day One developed?

The driver for Day One was a review of police initial recruitment in 2016, which highlighted concerns with the current Police SEARCH® recruitment model, including:

- a perceived lack of flexibility of the process
- a need for forces to run additional post-assessment centre processes to include the assessment of candidates' values and organisational fit
- the use of dated technology and limited changes in exercise format
- a perception of the process being cold and unwelcoming to candidates
- the process causing some contextual confusion as it is set in a shopping centre

It was recommended that the College of Policing support the MPS and a third-party consultancy to design and pilot a new assessment centre. The College of Policing provided subject-matter expert input throughout the assessment centre design process. Additional forces supported the design and latter live piloting including Surrey, Sussex, South Wales, Thames Valley, West Midlands and West Yorkshire.

During development, exercises were piloted and validated with serving police constables to ensure the exercises elicited responses relevant to effective performance in the rank of police constable.

To address the concerns raised regarding Police SEARCH®, the following design features have been put in place for Day One:

- During the design of Day One, care has been taken to maximise outcomes for underrepresented groups. More information on how Day One has integrated diversity by design into the process can be found in **Appendix A**.
- It is set within a policing context which lends itself to be a realistic job preview and improves the face validity of the assessment.
- Use of innovative technology which aids the assessment centre in being more efficient and provides a more modern experience for candidates and assessors.
- Day One is a competency and values-based assessment centre which uses the College of Policing's competency framework, known as the **Competency and Values Framework for Policing**.
- Assessors and assessment centre support staff have been trained to deliver the assessment centre in a manner which demonstrates warm and friendly behaviour towards candidates.

An assessor and facilitator course has been developed and implemented for the Day One Assessment Centre to equip assessors with the knowledge and skills needed to provide standardised and consistent assessing.

The College of Policing has developed guidance documents for the quality assurance of Day One and recommends assessors are quality assessed on a regular basis. Assessors are required to refresh their training so that assessing is conducted accurately and consistently.

Each candidate was required to undertake the same exercise suite. The exercise suite was comprised of: two role-play exercises (dispute settlement and early intervention), a structured interview, written exercise, virtual reality exercise and two video-based scenario exercises (burglary and youth engagement).

The remainder of this report provides information collated from the pilot Day One Assessment Centre.

2 Overall analysis

The following tables provide summary analysis for all candidates who attended a Day One assessment centre at each of the forces participating in the pilot.

Overall assessment centre success rate

	Attendance		Outcome			
	Total		Fail		Pass	
	Count	% (total)	Count	% (fail)	Count	% (pass)
All candidates	18,150	100%	6,230	34.3%	11,920	65.7%

Table 1 – Overall assessment centre success rate

Attendance and success rate by force

	Attendance		Outcome			
	Total		Fail		Pass	
	Count	% (total)	Count	% (fail)	Count	% (pass)
Metropolitan Police Service (MPS)	17,363	95.7%	6,079	35%	11,284	65%
Thames Valley Police (TVP)	518	2.9%	123	23.7%	395	76.3%
Surrey Police	█	1.1%	█	13.1%	█	86.9%
Sussex Police	█	0.4%	█	2.9%	█	97.1%

Table 2 – Overall assessment centre success rate by force

Attendance and success rate by sex

	Attendance		Outcome			
	Total		Fail		Pass	
	Count	% (total)	Count	% (fail)	Count	% (pass)
Female	6,414	35.3%	1,923	30%	4,491	70%
Male	11,535	63.6%	4,249	36.8%	7,286	63.2%
Prefer not to disclose	201	1.1%	58	28.9%	143	71.1%

Table 3 – Attendance and success rate by sex

Attendance and success rate by first language

	Attendance		Outcome			
	Total		Fail		Pass	
	Count	% (total)	Count	% (fail)	Count	% (pass)
English first language	10,518	58%	3,031	28.8%	7,487	71.2%
English not first language	3,119	17.2%	2,207	70.8%	912	29.2%
Not provided	4,513	24.9%	992	22%	3,521	78%

Table 4 – Attendance and success rate by first language

Attendance and success rate by disability

	Attendance		Outcome			
	Total		Fail		Pass	
	Count	% (total)	Count	% (fail)	Count	% (pass)
No disability	17,751	97.8%	6,115	34.4%	11,636	65.6%
Disability	399	2.2%	115	28.8%	284	71.2%

Table 5 – Attendance and success rate by disability

Attendance and success rate by age (banded)

	Attendance		Outcome			
	Total		Fail		Pass	
	Count	% (total)	Count	% (fail)	Count	% (pass)
18-24 years	8,043	44.3%	2,358	29.3%	5,685	70.7%
25-35 years	6,901	38%	2,295	33.3%	4,606	66.7%
36-60 years	2,986	16.5%	1,514	50.7%	1,472	49.3%
Not provided	220	1.2%	63	28.6%	157	71.4%

Table 6 – Attendance and success rate by age (banded)

Attendance and success rate by ethnicity (2+1)

	Attendance		Outcome				
	Total		Fail		Pass		AIR1
	Count	% (total)	Count	% (fail)	Count	% (pass)	
White (incl. White other)	12,300	67.8%	3,233	26.3%	9,067	73.7%	1.00
BAME	5,532	30.5%	2,895	52.3%	2,637	47.7%	0.65
Not stated	318	1.8%	102	32.1%	216	67.9%	0.92

Table 7 – Attendance and success rate by ethnicity (2+1) with AIRs¹

¹ AIR reference group: White (incl. White other)

Attendance and success rate by ethnicity (5+1)

	Attendance		Outcome				
	Total		Fail		Pass		AIR2
	Count	% (total)	Count	% (fail)	Count	% (pass)	
White (excl. White other)	10,131	55.8%	1,915	18.9%	8,216	81.1%	1.00
Black	1,408	7.8%	784	55.7%	624	44.3%	0.55
Asian	2,452	13.5%	1,426	58.2%	1,026	41.8%	0.52
Chinese	78	0.4%	25	32.1%	53	67.9%	0.84
Mixed	936	5.2%	283	30.2%	653	69.8%	0.86
Other	658	3.6%	377	57.3%	281	42.7%	0.53
White other	2,169	12%	1,318	60.8%	851	39.2%	0.48
Not stated	318	1.8%	102	32.1%	216	67.9%	0.84

Table 8 – Attendance and success rate by ethnicity (5+1) with AIRs²

² AIR reference group: White (excl. White other)

Attendance and success rate by ethnicity

	Attendance		Outcome					
	Total		Fail		Pass		AIR3	Age
	Count	% (total)	Count	% (fail)	Count	% (pass)		Mean (years)
Bangladeshi	417	2.3%	267	64%	150	36%	0.44	29.3
Black - Asian	35	0.2%	23	65.7%	12	34.3%	0.42	35.5
Black - African	532	2.9%	337	63.3%	195	36.7%	0.45	34.6
Black - British	325	1.8%	147	45.2%	178	54.8%	0.67	31.5
Black - Caribbean	179	1%	89	49.7%	90	50.3%	0.62	32.2
Black - Other	36	0.2%	21	58.3%	15	41.7%	0.51	32.8
Chinese	65	0.4%	20	30.8%	45	69.2%	0.85	29.4
Greek and Greek Cypriot	85	0.5%	48	56.5%	37	43.5%	0.54	30.8
Indian	551	3%	293	53.2%	258	46.8%	0.58	31.0
Mixed - White and Asian	229	1.3%	53	23.1%	176	76.9%	0.95	26.6
Mixed - White and Black African	99	0.5%	49	49.5%	50	50.5%	0.62	30.0
Mixed - White and Black Caribbean	208	1.1%	48	23.1%	160	76.9%	0.95	26.2

	Attendance		Outcome					AIR3	Age
	Total		Fail		Pass				
	Count	% (total)	Count	% (fail)	Count	% (pass)	Mean (years)		
Mixed - Other	210	1.2%	70	33.3%	140	66.7%	0.82	29.4	
Pakistani	543	3%	331	61%	212	39%	0.48	29.7	
Turkish and Turkish Cypriot	223	1.2%	137	61.4%	86	38.6%	0.47	27.7	
White - British	8,827	48.6%	1,651	18.7%	7,176	81.3%	1.00	26.7	
White - Irish	130	0.7%	31	23.8%	99	76.2%	0.94	29.7	
White - Other	1,726	9.5%	1,056	61.2%	670	38.8%	0.48	32.1	
Asian - Other	370	2%	223	60.3%	147	39.7%	0.49	31.5	
Other	235	1.3%	130	55.3%	105	44.7%	0.55	31.2	
White Welsh	■	0.1%	■	25%	■	75%	0.92	26.9	
White Scottish	■	0.0%	■	16.7%	■	83.3%	1.03	28.9	
Arabic	■	0.0%	■	100%	■	0%	0.00	50.8	
Not stated	3,106	17.1%	1,201	38.7%	1,905	61.3%	0.75	29.3	

Table 9 – Attendance and success rate by ethnicity with AIRs³

³ AIR reference group: White - British

Attendance and success rates by ethnicity (English as first language only)

	Attendance		Outcome					
	Total		Fail		Pass		AIR4	Age Mean (years)
	Count	% (total)	Count	% (fail)	Count	% (pass)		
Bangladeshi	169	1.6%	104	61.5%	65	38.5%	0.49	27.6
Black - Asian	20	0.2%	11	55%	9	45%	0.57	33.7
Black - African	209	2%	136	65.1%	73	34.9%	0.44	34.9
Black - British	196	1.9%	96	49%	100	51%	0.65	30.8
Black - Caribbean	132	1.3%	78	59.1%	54	40.9%	0.52	31.9
Black - Other	█	0.1%	█	62.5%	█	37.5%	0.47	24.1
Chinese	24	0.2%	6	25%	18	75%	0.95	28.0
Greek and Greek Cypriot	28	0.3%	10	35.7%	18	64.3%	0.81	30.0
Indian	223	2.1%	116	52%	107	48%	0.61	29.6
Mixed - White and Asian	148	1.4%	36	24.3%	112	75.7%	0.96	26.1
Mixed - White and Black African	59	0.6%	26	44.1%	33	55.9%	0.71	30.0
Mixed - White and Black Caribbean	149	1.4%	40	26.8%	109	73.2%	0.93	26.3

	Attendance		Outcome					
	Total		Fail		Pass		AIR4	Age
	Count	% (total)	Count	% (fail)	Count	% (pass)		Mean (years)
Mixed - Other	104	1%	33	31.7%	71	68.3%	0.86	28.1
Pakistani	234	2.2%	132	56.4%	102	43.6%	0.55	27.6
Turkish and Turkish Cypriot	86	0.8%	51	59.3%	35	40.7%	0.52	25.8
White - British	6,556	62.3%	1,378	21%	5,178	79%	1.00	26.6
White - Irish	96	0.9%	27	28.1%	69	71.9%	0.91	30.1
White - Other	266	2.5%	145	54.5%	121	45.5%	0.58	31.0
Asian - Other	118	1.1%	61	51.7%	57	48.3%	0.61	30.2
Other	66	0.6%	30	45.5%	36	54.5%	0.69	28.4
White Welsh	■	0.1%	■	25%	■	75%	0.95	26.9
White Scottish	■	0.0%	■	25%	■	75%	0.95	30.1
Not stated	1,611	15.3%	506	31.4%	1,105	68.6%	0.87	28.2

Table 10 – Attendance and success rate by ethnicity with AIRs (English as first language only)⁴

⁴ AIR reference group: White - British.

3 Exploration of biodata interactions

This section looks at interactions between ethnicity and other biographical factors. Ethnicity is explored in more detail, as this is the key factor (of all the protected characteristics reported) which shows lower success rates for underrepresented candidate groups.

Ethnicity representation of overall candidate pool and successful candidates

	White		BAME		Not stated	
	Count	% (total)	Count	% (total)	Count	% (total)
All candidates attending	12,300	67.8%	5,532	30.5%	318	1.8%
All successful candidates	9,067	76.1%	2,637	22.1%	216	1.8%

Table 11 – Ethnicity representation of overall candidate pool and successful candidates

Ethnicity by age

	Attendance		Ethnicity					
	Total		White		BAME		Not stated	
	Count	% (total)	Count	% (of White)	Count	% (of BAME)	Count	% (of not stated)
18-24 years	8,043	44.3%	6,019	48.9%	1,962	35.5%	62	19.5%
25-35 years	6,901	38%	4,596	37.4%	2,216	40.1%	89	28%
36-60 years	2,986	16.5%	1,623	13.2%	1,320	23.9%	43	13.5%
Not stated	220	1.2%	62	0.5%	34	0.6%	124	39%

Table 12 – Candidates' ethnicity (2+1) by age

Ethnicity by sex

	Attendance		Ethnicity					
	Total		White		BAME		Not stated	
	Count	% (total)	Count	% (of White)	Count	% (of BAME)	Count	% (of not stated)
Female	6,414	35.3%	4,586	37.3%	1,784	32.2%	44	13.8%
Male	11,535	63.6%	7,671	62.4%	3,731	67.4%	133	41.8%
Prefer not to disclose	201	1.1%	43	0.3%	17	0.3%	141	44.3%

Table 13 – Candidates' ethnicity (2+1) by sex

Ethnicity by first language

	Attendance		Ethnicity					
	Total		White		BAME		Not stated	
	Count	% (total)	Count	% (of White)	Count	% (of BAME)	Count	% (of not stated)
English first language	10,518	58%	7,860	63.9%	2,567	46.4%	91	28.6%
English not first language	3,119	17.2%	1,527	12.4%	1,545	27.9%	47	14.8%
Not stated	4,513	24.9%	2,913	23.7%	1,420	25.7%	180	56.6%

Table 14 – Candidates' ethnicity (2+1) by first language

Candidate success by ethnicity by force⁵

	White						BAME						Not stated		
	Fail		Pass		AIR	Fail	Fail		Pass		AIR	Fail	Pass	Fail	Pass
	Count	% (fail)	Count	% (pass)			Count	% (fail)	Count	% (pass)					
Metropolitan Police Service	3,111	26.7%	8,526	73.3%		2,875	52.6%	2,595	47.4%	0.65	93	36.3%	163	63.7%	
Surrey Police	20	14.7%	116	85.3%			12.5%		87.5%	1.03		9.3%		90.7%	
Sussex Police		3.1%		96.9%		0	0%		100%	1.03	0	0%	0	0%	
Thames Valley Police	100	21.6%	362	78.4%		19	38.8%	30	61.2%	0.78		57.1%		42.9%	
West Midlands Police	0	0%	0	0%		0	0%	0	0%	n/a	0	0%		100%	

Table 15 – Overall assessment centre success rate by ethnicity (2+1) by force

⁵ AIR reference group: White (within force).

4 Exercise performance

Cohen's *d* has been used to determine the standardised difference between two means (eg, BAME mean vs. White mean) and used to determine an appropriate effect size (the size of the difference between sets of means). The effect size definitions for this analysis have been taken from: Murphy, K. and Jacobs, R. (2012). Using effect size measures to reform the determination of adverse impact in equal employment litigation. *Psychology, Public Policy, and Law*, 18(3), 477–499.

Exercise analysis by ethnicity (low technology pilot forces only – Thames Valley Police, Surrey and Sussex)

Exercise	Ethnicity	N	Mean % score	SD	Min % score	Max % score	SEM	T	Sig	Mean	Pooled SD	Cohen's d	Effect size (see Appendix E)
Early Intervention	White	663	57.3	13.0	20.0	95.0	0.50	0.3	n.s	0.6	13.46	0.04	Small
	BAME	62	56.8	13.9	25.0	80.0	1.77						
Dispute settlement	White	663	59.1	13.1	20.0	95.0	0.51	-0.5	n.s	-0.8	12.91	-0.06	Small
	BAME	62	59.8	12.7	25.0	90.0	1.61						
Stop and search	White	663	56.3	12.4	20.0	86.7	0.48	1.8	n.s	2.8	12.18	0.23	Small to medium
	BAME	62	53.4	12.0	33.3	80.0	1.25						
Collaboration	White	663	58.0	15.5	20.0	100.0	0.60	3.7	0.001	7.8	15.60	0.50	Medium
	BAME	62	50.2	15.7	20.0	80.0	2.00						
Interview	White	663	63.6	13.1	20.0	95.0	0.51	-1.1	n.s	-2.2	13.91	-0.16	Small
	BAME	62	65.7	14.7	25.0	95.0	1.86						
Youth engagement	White	663	69.1	12.7	20.0	100.0	0.49	1.6	n.s	2.7	12.84	0.21	Small to medium
	BAME	62	66.4	13.0	35.0	95.0	1.65						
Burglary	White	663	79.6	13.5	33.3	100.0	0.53	2.6	n.s	5.8	15.25	0.38	Medium
	BAME	62	73.8	16.8	33.3	100.0	2.13						

Table 16 – Exercise analysis by ethnicity (low technology pilot forces only – Thames Valley Police, Surrey and Sussex)

Exercise analysis by ethnicity

Exercise	Ethnicity	N	Mean % score	SD	Min % score	Max % score	SEM	T	Sig	Mean	Pooled SD	Cohen's d	Effect size (see Appendix E)
Early Intervention	White	12,300	55.3	15.1	20.0	100.0	0.14	22.7	0.001	5.9	15.8	0.37	Medium
	BAME	5,532	49.4	16.5	20.0	100.0	0.22						
Dispute settlement	White	12,300	59.5	14.7	20.0	100.0	0.13	23.7	0.001	6.0	15.4	0.39	Medium
	BAME	5,532	53.5	16.0	20.0	100.0	0.21						
Stop and search	White	12,300	53.3	14.5	20.0	100.0	0.13	28.5	0.001	6.7	14.5	0.46	Medium
	BAME	5,532	46.6	14.5	20.0	100.0	0.19						
Collaboration	White	12,300	59.9	15.2	20.0	100.0	0.14	30.7	0.001	7.9	15.7	0.50	Medium
	BAME	5,532	52.0	16.2	20.0	100.0	0.22						
Interview	White	12,300	65.0	13.1	20.0	100.0	0.12	16.0	0.001	3.6	13.5	0.26	Small to medium
	BAME	5,532	61.4	14.0	20.0	100.0	0.19						
Youth engagement	White	12,300	66.6	13.2	20.0	100.0	0.12	19.9	0.001	4.5	13.6	0.33	Small to medium
	BAME	5,532	62.1	14.1	20.0	100.0	0.19						
Burglary	White	12,300	74.0	15.3	20.0	100.0	0.14	38.5	0.001	10.4	16.3	0.64	Medium
	BAME	5,532	63.7	17.2	20.0	100.0	0.23						

Table 17 – Exercise analysis by ethnicity

Exercise analysis by sex

Exercise	Sex	N	Mean % score	SD	Min % score	Max % score	SEM	T	Sig	Mean	Pooled SD	Cohen's d	Effect size
Early Intervention	Female	6,414	54.1	15.3	20.0	100.0	0.19	4.2	0.001	1.0	15.7	0.06	Small
	Male	11,535	53.1	16.1	20.0	100.0	0.15						
Dispute settlement	Female	6,141	58.5	15.1	20.0	100.0	0.19	5.5	0.001	1.3	15.3	0.09	Small
	Male	11,535	57.2	15.5	20.0	100.0	0.14						
Stop and search	Female	6,414	53.6	14.8	20.0	100.0	0.19	16.2	0.001	3.7	14.8	0.25	Small to medium
	Male	11,535	49.8	14.7	20.0	100.0	0.14						
Collaboration	Female	6,414	57.6	15.9	20.0	100.0	0.20	1.0	n.s	0.2	16.0	0.02	Small
	Male	11,536	57.3	16.0	20.0	100.0	0.15						
Interview	Female	6,414	65.5	13.1	20.0	100.0	0.16	12.3	0.001	2.6	13.4	0.19	Small
	Male	11,535	63.0	13.6	20.0	100.0	0.13						
Youth engagement	Female	6,414	65.9	13.8	20.0	100.0	0.17	5.1	0.001	1.1	13.7	0.08	Small
	Male	11,535	64.8	13.5	20.0	100.0	0.13						
Burglary	Female	6,414	72.2	16.4	20.0	100.0	0.21	8.7	0.001	2.2	16.5	0.13	Small
	Male	11,535	70.0	16.6	20.0	100.0	0.15						

Table 18 – Exercise analysis by sex

5 Discussion

This section highlights key differences in candidates' assessment outcomes and sets out a summary of action taken by the College ahead of the national rollout of the new Day One Assessment Centre in the second half of 2020.

Before key differences are discussed, it is worth highlighting that there were a number of underrepresented groups who performed better than their comparative majority group. For example, female candidates were more successful than male candidates (70% vs. 63.2%), and candidates who reported a disability had a higher success rate than those without a disability (71.2% vs. 65.6%).

One of the key questions about any selection method is whether it has 'adverse impact' against any groups with protected characteristics under the Equality Act. Adverse impact in the context of an assessment is used to describe assessment processes that have a differential and negative impact on candidates from minority or protected characteristics groups.

Adverse impact ratio (AIR) – Adverse impact is measured by means of a ratio (AIR). The ratio is calculated by dividing the pass rate of the minority group (or the protected characteristic group) by the pass rate of the majority group. This results in a figure that varies between 0 and 1. Generally the assessment industry applies the '0.8 or four-fifths rule' – if the AIR is less than 0.8 or four fifths, when comparing two groups, this may provide evidence of adverse impact, and the organisation is required to investigate the causes of the differences. An AIR below 0.8 is not always an indicator of faults or issues with the selection measure itself, as other factors such as attraction strategy, ability of candidates, and presence or absence of positive action can affect this.

Ethnicity – there are differences in pass rates between White and BAME candidates (with BAME candidates achieving lower success rates). Some minority ethnic groups have pass rates which generate AIRs above 0.8 (eg, Chinese = 0.85, White and mixed Asian = 0.95, and White and Black Caribbean = 0.95). However, some candidate subgroups have AIRs lower than 0.8 (e.g. Black Asian = 0.42, Bangladeshi = 0.44, Pakistani = 0.48). Possible reasons for these differences are discussed in this section, which include the influence of other biographical factors and differences in pre-selection processes across different pilot forces.

A number of interactions were identified between ethnicity and other biographical factors. Analysis to date indicates the following general findings:

- **Age:** BAME candidates tended to be older than White candidates (BAME candidates had an average age of 30.5 years, compared to 27.6 years for White candidates).
- **Sex:** a higher proportion of BAME candidates were male.
- **Second language:** significantly higher proportions of BAME candidates had English as a second language. Whether or not the candidates have English as a first language does create a difference in assessment outcomes – particularly for the MPS candidate pool. Based on research conducted in 2019⁶ the level of English language required at the assessment centre is no higher than Functional Skills Level 2, which is appropriate for police recruitment. An overview of Functional Skills Level 2 English Language is set out in **Appendix B**.

Force differences – There are clear variations in overall success rates across forces (MPS = 65%, TVP = 76.3%, Surrey = 86.9% and Sussex = 97.1%). It is likely that the difference in success rates is linked to differences in pre-assessment processes and different candidate pools across forces. For example, the MPS did not apply any pre-selection sift during the pilot period – all candidates were invited to the assessment centre. Based on research conducted in 2019, it is likely a high proportion of the MPS candidate pool do not possess Functional Skills Level 2 in English Language. These factors are likely to have affected the BAME assessment outcomes.

The evidence suggests the gap in success rates between some BAME groups and the White group at Day One will be smaller for forces who implement appropriate pre-selection or sifting of candidates. The data/results from Surrey and Sussex support this. They have introduced positive action and sifting into their recruitment processes and this has resulted in positive outcomes in terms of proportions of successful BAME candidates (eg, Surrey has a 100% success rate for BAME candidates and

⁶ In August 2019, 187 MPS candidates who attended the Day One Assessment Centre were also assessed for Functional Skills Level 2 English Language (oral and written). 59.9% (112) of the sample met the standard for Day One. 34% of the sample met the standard for Functional Skills Level 2 English Language.

Sussex has an 87.5% success rate for BAME candidates). While caution is required in extrapolating these findings as the candidate numbers are low, these are encouraging findings. The evidence suggests that implementing an appropriate recruitment strategy, including investment in attraction, sifting and positive action initiatives may be an effective means of tackling subgroup differentials at the Day One Assessment Centre. However, forces need to ensure that adverse impact is not being transferred to earlier stages of the recruitment process.

Exercise performance – there were performance differences between BAME and White candidates across the exercise suite. There were variations in size of difference between the pilot centres (see tables 17 and 18). These differences were small to medium. There were no exercises where large differences were found between the percentage scores of BAME and White candidates undertaking the pilot at TVP or the MPS. The collaboration and burglary exercises had the largest differences within each of the pilot centres. These differences were generally larger at the MPS centre.

The College is reviewing the exercise suite to take account of where differences in candidate subgroups exist. Actions to be taken ahead of rollout are set out towards the end of this section.

Further relevant information

- **Comparison with constables across the service** – data reported by the statistical bulletin (up to 31 March 2019) indicates that 7.3% of constables nationally are from a BAME background (16% of constables in the MPS are BAME) and 33.4% are female (30.1% of constables in the MPS are female). In comparison, the proportion of Day One candidates from a BAME background is 30.5% and the proportion of female Day One candidates is 35.3%. The proportion of successful BAME candidates at Day One is 22.1% and the proportion of successful female candidates is 37.7%.

- **Candidate feedback**⁷ – during the pilot all candidates were asked to complete a candidate survey. 86.71% of Day One candidates who completed a survey after their assessment ‘agreed’ or ‘strongly agreed’ that they were satisfied with delivery of the assessment centre. Day One candidates indicated through their survey responses that they were more likely to say they would recommend the assessment centre to a friend or colleague than SEARCH[®] candidates (85% of Day One candidates vs. 79% of SEARCH[®] candidates), and this effect was more pronounced for BAME candidates (64% of BAME Day One candidates vs. 39% of BAME SEARCH[®] candidates).

While analysis in this report gives some initial insights into possible factors influencing biographical differences between groups, further exploration, particularly using multivariate statistical analysis, will establish which biographical factors (including ethnicity) are most associated with success at the Day One Assessment Centre.

Action ahead of national rollout – below sets out action that will be taken by the College ahead of rollout based on data analysed to date:

- **Final exercise suite:** removal of collaboration exercise in the rollout of Day One. The collaboration exercise will be removed as there is evidence across two pilot centres that this exercise generates larger subgroup differences, with the BAME subgroup overall scoring lower than the White group. There is a small but significant difference in performance between candidates in the two candidate roles⁸.
- **Burglary exercise** – as the difference between BAME and White candidates at the MPS for the burglary exercise is high-medium, the College is reviewing exercise content with diversity and operational subject matter experts to identify possible issues within the content and opportunities to further enhance the exercise. This work will include piloting a refined version of the exercise with a diverse candidate audience. An updated version of the exercise will be included in the final exercise suite.

⁷ Based on the completion of 3,829 candidate surveys between 4 April 2018 and 31 August 2019.

⁸ Within the collaboration exercise, one candidate takes on the role of a tablet device user and the other candidate takes on the role of a virtual reality headset user.

- **Refinement of written exercise:** consistency of assessors has been measured through an inter-rater reliability study (reported within the Day One interim evaluation report). Evidence indicates that assessors would benefit from further support to assist them assessing the written exercise. The written exercise content and assessor guidance materials are currently being refined.
- **Candidate management guidance:** the College is working with the Police Uplift Programme on candidate management. This includes a national attraction campaign targeting underrepresented groups, providing guidance and support to forces on candidate management (which will include guidance on attraction, sifting and positive action). In the medium term, commercial approval will be sought for the procurement of an appropriate national sift mechanism for forces to use.
- **Change to exercise checklists** – during piloting, evidence was collated from assessors which indicated that the Observe, Record, Classify and Evaluate (ORCE) approach to assessing was presenting challenges with the volume of candidates required to be assessed. To improve assessor experience and consistency in assessing, the final suite of exercises will include checklists for the two role plays and the written exercise, while the ORCE method will remain for the interview.

Day One will be continually reviewed and improved in the same way as all College assessment processes are, in line with industry standards and best practice.

6 Next steps

Day One was due to be rolled out to all forces from the summer of 2020. However, the coronavirus pandemic and subsequent national lockdown from March 2020 meant that all face-to-face officer recruitment was postponed. The College of Policing was tasked to develop and deliver an entirely online recruitment process at pace, to enable forces to continue to deliver on the government's commitment to recruit 20,000 additional police officers over the next three years.

The online assessment process has been successfully implemented by the College and initial findings are that the process is working well. The College will continue to monitor both the effectiveness of the process and the outcomes in relation to diversity, and will conduct a formal evaluation in the autumn of 2020.

The online assessment process will remain in place throughout 2020, to ensure consistency and fairness to candidates throughout the period of the pandemic. The College will review the evidence in relation to Day One, as well as the evaluation of the online assessment process, before setting out plans for the longer-term future of police assessment. These plans may incorporate elements of online and face-to-face assessment.

This evidence will include the outcomes of an independent review of the Day One assessment centre that was conducted in March 2020. The review concluded that, overall, the College has produced a reliable, valid and objective set of selection exercises incorporating many aspects of best practice. The review also found that the exercises provide clear guidance for assessors, quality assurers and candidates. The review set out a number of recommendations, which the College will address ahead of any further implementation of the Day One exercises.

The College is developing a publication strategy for future release of data relating to the national recruit assessment products.

7 Logistical regression analysis

A Chartered Occupational Psychologist used a logistical regression analysis to explore which biographical factors were most associated with differences in success at the Day One assessment centre. There are a number of factors unaccounted for within the model which are likely to influence success at Day One (eg, educational attainment, level of English language attained and socio-economic status). However, the model was moderately successful in explaining variance in candidate success and provided a good statistical fit. See **Appendix C** for more detail about the methodology.

Data from 10,730 candidates was included in the analysis. The findings of the analysis are as follows:

- Candidates for whom English is a first language were more likely to pass the assessment centre than candidates for whom English is not their first language. However, this effect was less pronounced for BAME candidates than for White candidates (ie, difference in success rates was smaller for BAME candidates regardless of whether they have English as a first language or not) – see figure 1 in **Appendix D**.
- For candidates for whom English is a first language, male and female candidates performed similarly, but for candidates who do not have English as a first language, females were more likely to be successful – see figure 2 in **Appendix D**.
- Female candidates were more likely to pass than male candidates, and this is the case for both BAME and White candidates – see figure 3 in **Appendix D**.
- Candidates for whom English is a first language were more likely to be successful across all the religious background groups. However, the effect of having English as a first language was strongest for Christian candidates and weakest for Muslim candidates – see figure 4 in **Appendix D**.
- Candidates from non-Metropolitan Police Service (MPS) forces were more likely to pass than MPS candidates – this is explained by the significantly higher pass rate for non-MPS forces.

These results overall show that first language, religious background, force and sex were all more predictive of success at the Day One assessment centre than ethnicity on its own. However, the relationships between the different variables are complex.

While some biographical variables have a greater influence on assessment centre success, they do not fully explain the results on their own. The analysis, while providing insights, is limited by biographical variables unaccounted for (eg, academic attainment, level of English language attained, socio-economic status) in the model and is restricted to the four forces who participated in the pilot. In considering the current analysis, it is recommended that:

- Greater breadth of data needs to be collated on additional biographical data such as educational attainment, level of English language attained and socio-economic status. These need to be collated across policing in a consistent and coordinated way.
- Qualitative analysis is undertaken to understand why candidates from different subgroup backgrounds are performing less well than other groups.
- The regression analysis is repeated when more data has been collated across more forces.

Appendix A: Diversity by design

Based on the latest evidence base⁹ a number of actions were taken to maximise outcomes for underrepresented groups. These actions are summarised below.

Designed to measure the competency and values framework	Incorporates updated competencies and the values underpinning the Code of Ethics. Values-based recruitment potentially improves diversity outcomes, retention and wellbeing.
Use of exercises that maximise performance of underrepresented groups	Uses exercises known to have lower subgroup differences and removes exercises known to have higher subgroup differences based on best available evidence.
Designed to deliver a positive candidate experience	Using technology-driven exercises and interactive role plays delivers a modern and professional candidate experience and presents the service in a positive light.
Candidate experience warm and welcoming	Evidence suggests a positive candidate experience has a positive impact on performance of candidates from underrepresented groups.
Incorporates auto-scored exercises	Auto scoring eliminates any potential for assessor bias, where used.
Inclusion of underrepresented groups in the design process	Pre-live pilot and validation cohorts were 29% BAME and 58% female. Job evaluation sample was 33% BAME and 48% female.
Diverse bank of assessors	The MPS bank of assessors for Day One is 51% BAME and 58% female.

⁹ College of Policing review of police initial recruitment (2016)

Appendix B: Overview of Functional Skills Level 2 English Language

During the evaluation, a sample (187) of MPS candidates' Functional Skills Level 2 English Language competence was assessed for Speaking, Listening and Communication Level 2 and Writing Level 2. The detail of relevant Functional Skills Level 2 assessment criteria is set out below:

Speaking, Listening and Communication Level 2

- Make a range of contributions to discussions in a range of contexts, including those that are unfamiliar.
- Consider different sources of information and give a relevant, clear response.
- Present information and ideas clearly and persuasively to others.
- Adapt contributions to suit audience, purpose and situation.

Writing Level 2

- Present information/ideas concisely, logically and persuasively.
- Present information clearly and concisely.
- Use a range of writing styles for different purposes.
- Use a range of sentence structures, including paragraphs, to organise written communication effectively.
- Punctuate written text using commas, apostrophes and inverted commas accurately. Make sure written work is fit for purpose and audience, with accurate spelling and grammar that support clear meaning.

Appendix C: Logistical regression methodology

For the purposes of statistical modelling, some aggregation of biographical factors was necessary to conduct meaningful analysis. Where aggregation of categories occurred, consideration was given to the potential impact on the analysis. As further data are collected, there will be a greater opportunity to explore the impact of biographical variables on assessment centre results without aggregation of categories.

The biographical factors considered for the analysis were first language, ethnicity, age, force, sex, disability, sexual orientation and religion. Any missing values or 'prefer not to say' responses were removed from each of the variables, which resulted in a reduction in the sample size from 18,150 to 10,730. 85.2% of White candidates had English as a first language (6,643 out of 7,793), compared to 63.6% of BAME candidates (1,868 out of 2,937). The effects of first language will have a greater relative impact on the overall success rate of BAME candidates, due to the higher proportion of candidates for whom English is not their first language.

While some caution needs to be taken in interpreting the statistics, approximately 23% of the difference (variance) in assessment centre results was explained by the logistical regression model.

Other biographical factors (eg, age and disability) show moderate differentials in terms of success rate. The descriptive statistics show that younger candidates tended to have slightly higher success rates, as did candidates with a disability. While these terms were not fitted within the regression model on this occasion, it is important to continue to monitor their relationship with success rates within future analyses.

Appendix D: Interactions within logistical regression model

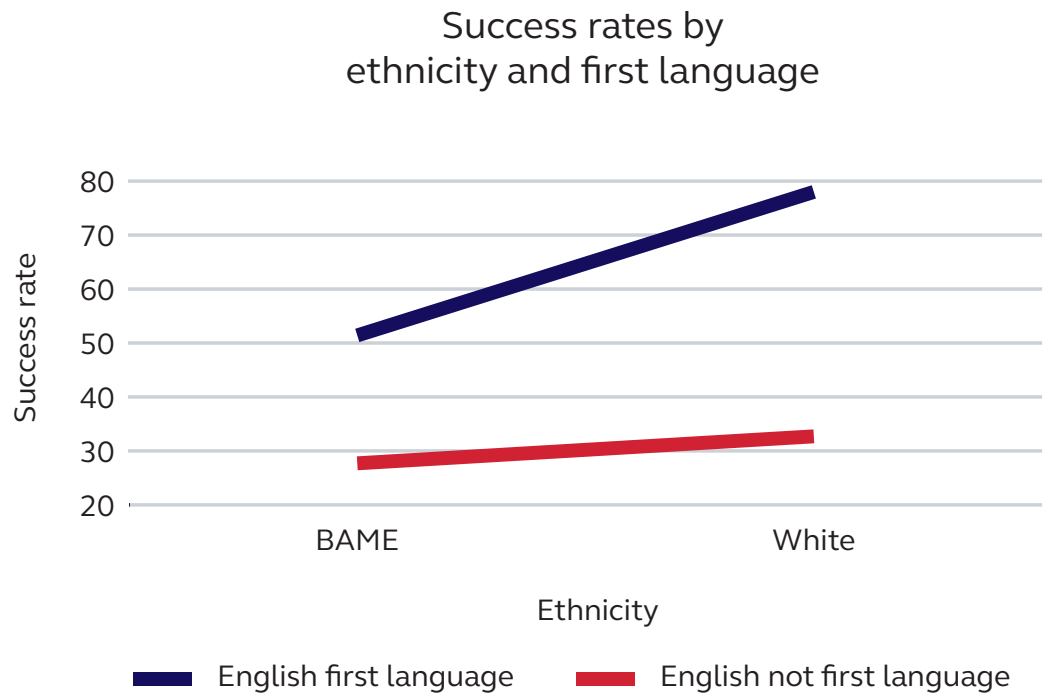


Figure 1 – Relative success rates by ethnicity and first language

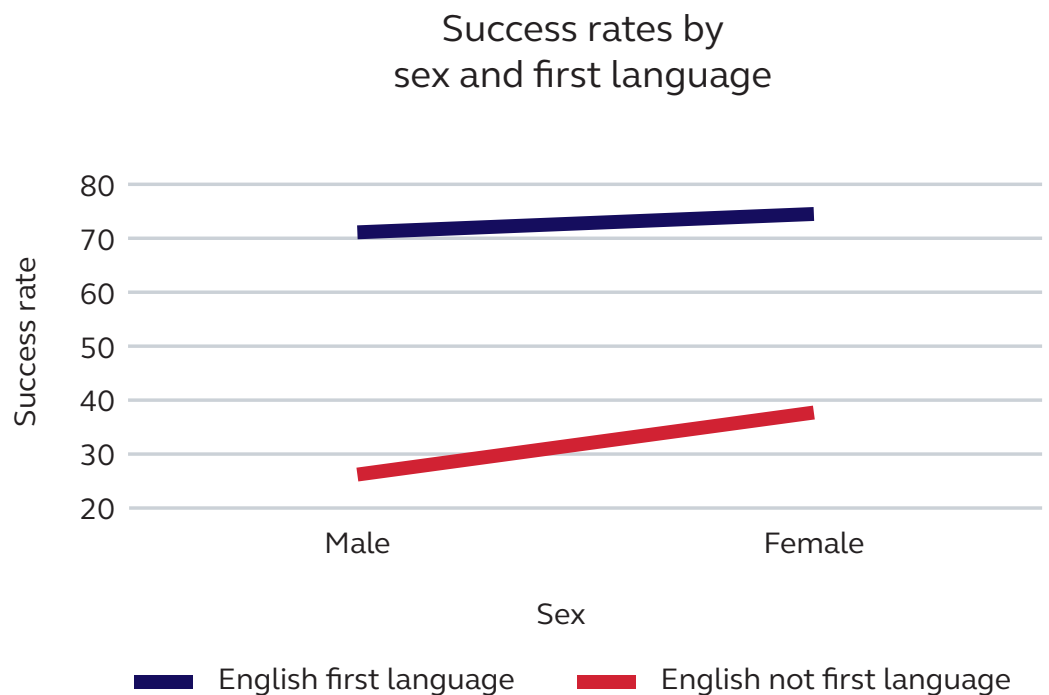


Figure 2 – Relative success rates by sex and first language

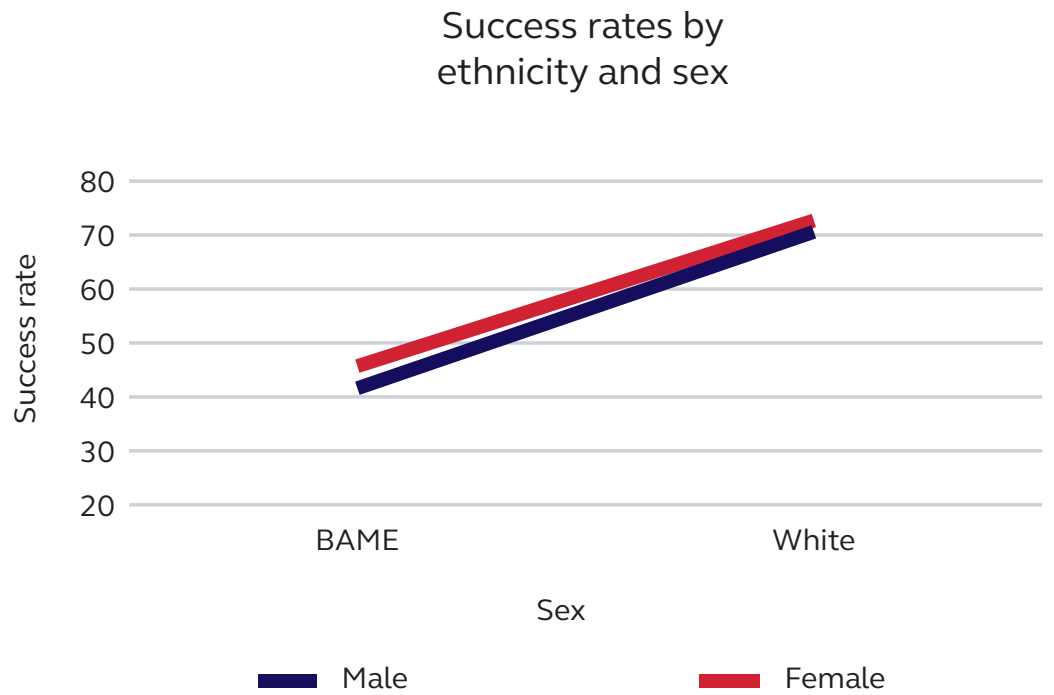


Figure 3 – Relative success rates by ethnicity and sex

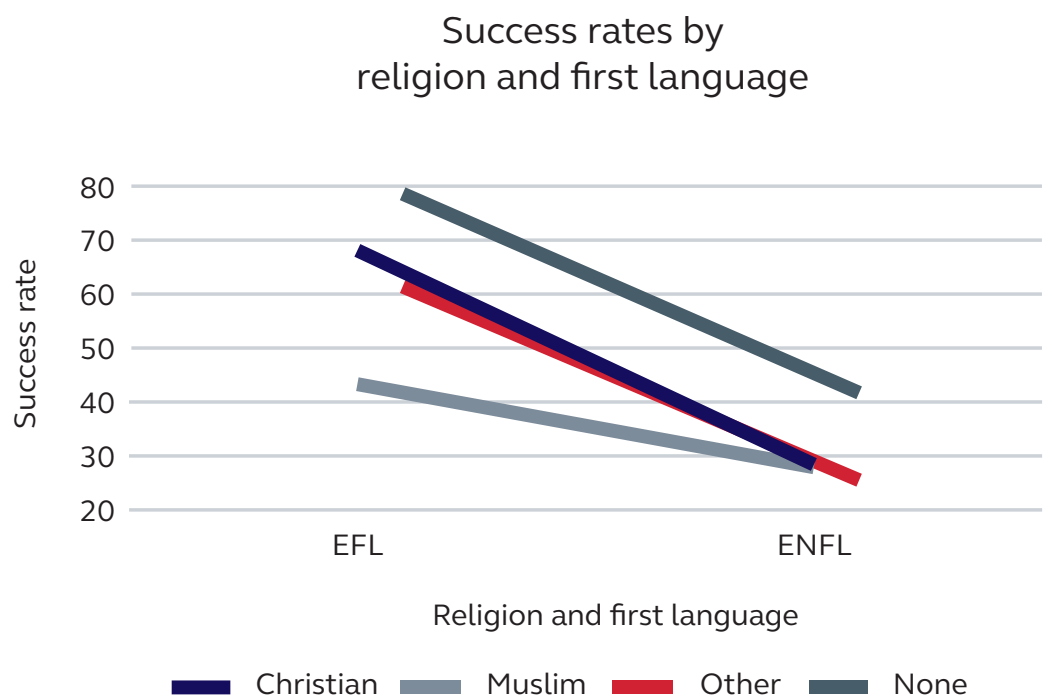


Figure 4 – Relative success rates by religion and first language

Appendix E: Effect size labels

The statistic used to measure effect size is Cohen's d. Cohen defines an effect size of 0.2 as small, 0.5 as medium and 0.8 as large.* As these reflect specific effect sizes as opposed to ranges, the following convention has been used to label effect sizes, in line with Cohen's guidelines:

- <0.2: small
- 0.2-0.34: small to medium
- 0.35-0.64: medium
- 0.65-0.79: medium to large
- ≥ 0.8 : large

* Cohen J. (1988). 'Statistical Power Analysis for the Behavioral Sciences'. 2nd ed. New York: Routledge.

About the College

We're the professional body for everyone who works for the police service in England and Wales. Our purpose is to provide those working in policing with the skills and knowledge necessary to prevent crime, protect the public and secure public trust.

college.police.uk

 Follow us
[@CollegeofPolice](https://twitter.com/CollegeofPolice)