Case No: 1304402/2020

EMPLOYMENT TRIBUNALS

Claimant: Mr A Girgis

Respondent: Joint Committee on Intercollegiate Examinations

Heard at: Midlands West

On: 3, 4, 6, 7, 10, 11, 12, 13, 14 (in chambers) and 17 July 2023 and (in chambers) 18 to 21 July 2023

Before: Employment Judge Faulkner
Mrs J Keene
Mr R Virdee

Representation: **Claimant** - Mr N Grundy (counsel)
Respondent - Mr M Briggs (counsel)

RESERVED JUDGMENT

1. The Respondent did not contravene section 53 of the Equality Act 2010 by:

1.1. Discriminating against the Claimant in the arrangements it made for deciding upon whom to confer a relevant qualification.

1.2. Discriminating against the Claimant by not conferring a relevant qualification on him.

2. The Claimant's complaints are dismissed accordingly.

3. The Remedy Hearing provisionally arranged for 4, 5 and 6 December 2023 is cancelled. The Case Management Orders made in relation to that Hearing are rescinded.

REASONS

1. The parties made clear early on in this Hearing that, whatever the outcome, both would require written reasons for it. The Reasons below are provided in response to that request.

Complaints

2. This Claim was concerned with complaints made pursuant to section 53 of the Equality Act 2010 ("the Act"), the Respondent being a "qualifications body" for the purposes of that section. In short, the Claimant complained of indirect discrimination in relation to the arrangements the Respondent made for deciding upon whom to confer a relevant qualification (section 53(1)(a)) and of direct discrimination in relation to the Respondent's decision not to confer that qualification on him (section 53(1)(c)).

Issues

3. The parties had produced an almost-agreed list of issues prior to the Hearing. It was agreed on day 1 that the issues to be decided were definitively as follows, augmented by the Claimant's successful amendment application on day 2 referred to below.

Direct discrimination

4. The first question to be decided in relation to direct discrimination was whether the Respondent, in not conferring a relevant qualification on the Claimant, treated him less favourably than it would have treated a UK national in marking his examination taken on 14 November 2019? The comparator, it was agreed, was someone who answered the examination questions in the same way as the Claimant. There was no complaint that the justifications/explanations given for the Claimant's marks were themselves also discriminatory.

5. If so, the second question was whether that was because of race. The Claimant relied on being a non-UK national, which Mr Briggs accepted was a satisfactory basis for the complaint. No case was put on the basis of the Claimant being Egyptian or Coptic, as Mr Grundy accepted in closing submissions.

Indirect discrimination

6. As to indirect discrimination, the first question to be determined was whether the Respondent applied to the Claimant one or more of the following provisions, criteria or practices ("PCPs") in relation to the examination conducted on 14 November 2019 (the Respondent accepted that it applied the first, third and fourth and that they were PCPs, but did not accept that it applied the second):

6.1. Requiring candidates to score 576 marks to pass the examination.

6.2. Marking down candidates who were not in UK/EU training posts and/or marking down non-UK national candidates and/or minority ethnic candidates.

6.3. In the format of the examination, testing whether candidates were suitable to be day 1 consultants in urology in the UK.

6.4. (Permitted by way of amendment on day 2) allowing candidates (who were fit and well) a maximum of four attempts to pass the examination.

7. The second question was whether any PCP was applied to all candidates for the examination. As will appear below, putting the question in that way when agreeing the list of issues did not determine the pool of candidates by which the third question below was to be assessed, but the Respondent accepted that it applied the first, third and fourth PCPs to all candidates.

8. The third question was whether any PCP put persons of Egyptian and/or non-UK nationality at a particular disadvantage compared to persons not of Egyptian and/or of UK nationality, in that they were less likely to pass the exam. In fact, as Mr Grundy conceded in closing, the Claimant's case was only put on the basis of UK/non-UK nationality throughout. The Respondent did not accept that any PCP put non-UK nationals at a particular disadvantage.

9. The fourth question, if so, was whether it put the Claimant at that disadvantage. The Respondent accepted that the first and third PCPs did.

10. If so, the final question was whether the Respondent had shown that the PCP(s) in question was a proportionate means of achieving a legitimate aim. It relied on the aim of needing to keep patients and the public safe and give them confidence in consultants.

Time limits

11. There were no time limit issues in respect of most of the complaints. The Claim Form was treated as received on 22 April 2020. Employment Judge Cookson extended time for presentation of the Claim at a preliminary hearing. The complaint about PCP4 however is deemed to have been presented when the Tribunal gave the Claimant permission to amend his Claim, namely on day 2 of this Hearing, 4 July 2023. If that complaint succeeded, the Tribunal was required to determine whether it was presented within such further time after expiry of the normal time limit as the Tribunal thinks just and equitable.

Amendment

12. On day 2, the Claimant sought to reformulate the fourth PCP in the terms set out above, relying on same particular disadvantage which he said flowed from it. The fourth PCP had originally read, "Providing candidates in UK training posts with additional tutoring and teaching days to prepare for the examination". The reformulation was not set out in writing, as case law strongly recommends, but no point was taken about that and of course we were at the Final Hearing where that is not always practicable. It was also a reformulation that was short and easy to understand.

13. The Claimant submitted that it was not a new point. It was not new in that both parties had always known that the Respondent had a regulation that meant when the Claimant failed the examination in question on the fourth occasion, he could not take it again. That was not however the same as saying that the

regulation/arrangement had previously been challenged as indirectly discriminatory. At Mr Grundy's direction we considered the following documents:

13.1. The Claim Form (pages 10 to 24) stated the fact that the Respondent allows a maximum of four attempts with no re-entry. What the Claimant – who we accept was unrepresented at the time – explicitly said was discriminatory however was the examiners marking him down and failing him and the examination being failed in disproportionately large numbers by people who were medically qualified outside the UK/EU. This essentially corresponds with his direct and indirect discrimination complaints respectively as set out in the list of issues above, and in either case it was the failing of the examination that was explicitly attacked as discriminatory and not the rule regarding maximum attempts. It was notable too that in his long list of what he sought if successful in his Claim (page 17), he did not say that the Respondent should consider more than four attempts being permitted. For all of these reasons we concluded that the Claimant was plainly not challenging the maximum attempts rule in the Claim Form.

13.2. In the document the Claimant prepared in reply to the Tribunal's request for comments (pages 55 to 61, dated 26 June 2020), the Claimant stated that none of the eleven candidates who failed in their four attempts were UK-trained or native British. That was plainly not however, when read in context, an attack on the rule about maximum attempts, but part of the Claimant's statistical evidence by which he sought to show that the requirement to pass the examination was discriminatory. The PCP was explicitly said to be the requirement to pass the examination, not the four attempts rule and in fact the statistical evidence he referred to explicitly attacked the pass rate not the number of attempts.

13.3. In the Claimant Further and Better Particulars, drafted by Mr Grundy (pages 71 to 82, undated) at page 78, the Claimant provided an answer to question 8 posed by the Respondent. That question asked for evidence that his marks were because of race i.e., that they were the result of direct discrimination. Moreover, whilst again the Claimant cited that 100% of UK nationals passed the fourth attempt but only 21% of non-UK nationals did so, that was yet again a plain reference to the passing of the exam itself (the pass rates) not the rule limiting candidates to four attempts. This conclusion and that at paragraph 13.2 above was confirmed in our view by Mr Grundy's statement in closing submissions that the Claimant relied on this statistical difference as his fallback evidence in support of his case that PCPs 1 and 3 put non-UK nationals at a particular disadvantage.

14. In our view therefore, a complaint of indirect discrimination resulting from PCP4 as reformulated was plainly not within the four walls of the existing case and the reformulation was thus to be treated as a substantive amendment application.

15. It was clearly a late application, something that was especially notable given that the Claimant had been professionally advised for over three years by the time the application was made, though we accepted Mr Grundy's submission that this was by no means a determinative factor. Case law from **Selkent Bus Co Ltd v Moore [1996] ICR 836** onwards, particularly recently in the Employment Appeal Tribunal ("EAT") in **Vaughan v Modality Partnership [2021] ICR 535**,

makes clear that the key factor in deciding amendment applications is the balance of prejudice.

16. There would have been prejudice to the Claimant in refusing the application because what he sought to add to his Claim was a different line of argument, as far as we could tell at what was for us an early stage in considering the case, compared to the other PCPs he was seeking to attack as discriminatory. As to whether there would be prejudice to the Respondent in allowing the application

16.1. It was of course aware of the rule, though that was not the main point for our consideration. It did not know until the second day of this Hearing that the Claimant was attacking it as discriminatory.

16.2. Mr Briggs accepted (once the point was clarified by Mr Grundy) that the Respondent had the evidence on which the Claimant relied to prove particular disadvantage flowing from the PCP, and had had it for some time, namely the percentages of those who failed the fourth exam amongst UK and non-UK nationals respectively.

16.3. The Respondent's witnesses who were examiners in November 2019 had not in their statements addressed the PCP, its alleged discriminatory effect or its potential justification, though we noted that Mr Briggs' submissions opposing the application focused solely on concerns about whether the Respondent could put forward a justification defence once Mr Grundy had made clear the basis of the Claimant's case of particular disadvantage.

16.4. Mr Briggs stated that the legitimate aim would be the same as that already relied on in respect of the other PCPs. Up to that point therefore, there was no identifiable prejudice to the Respondent because it was or should have been ready to speak to the legitimate aim, and in any event to our mind it was hardly arguable that the aim was legitimate, as the Claimant eventually conceded.

16.5. The crucial question was proportionality. We accepted that this was not dealt with in the Respondent's statements. As things stood at the time of deciding the application however, it seemed to us that the point was a discrete one and not complex to consider. It seemed apparent that the relevant witnesses – Mr Ryan and possibly Mr Dickinson – should be able to deal with it either in examination in chief or in supplementary statements (we were inclined to think the former would be sufficient). Whilst we accepted that both witnesses were busy medical practitioners, there was time in the Hearing – the following day was a non-sitting day and indeed there were several more days before they would be giving evidence – for one or both of them to prepare to deal with the point.

17. The merits of a complaint based on a proposed amendment can be taken into account – **Kumari v Greater Manchester Mental Health NHS Foundation Trust [2022] EAT 132** – and Mr Briggs did seem in part to resist the amendment on this basis. If the PCP as pleaded proved to be an unsustainable case for the Claimant, that was of course a matter for him, and we would certainly have taken some persuading to allow even later amendments than this one, but we were not in a position on the morning of day 2 of the Hearing to assess the merits of the complaint. **Kumari** makes clear that a tribunal should only take the merits into account if a proper assessment of them can be made.

18. Weighing up the parties' respective positions, we allowed the amendment – and made clear it was an amendment. This was strictly subject to time limit issues, which we were not in a position to decide at this stage, not least because we had not heard evidence from the Claimant as to the reason for the delay and its length. **Galilee v Commissioner of Police of the Metropolis UKEAT/0207/16** says that time limit issues can be parked over to be determined at a final hearing. We were at that stage when allowing the amendment, but determined – as both counsel in submissions on the matter indicated was the correct course of action – that this should be held over to our final determination.

19. As set out in the list of issues above therefore, the question was whether the time limit should be extended, the complaint being deemed to be presented on 4 July 2023. Prejudice is one of the factors that should be considered in dealing with that question and of course our position on that was relevant to deciding the amendment application. That said, first, we made clear to the parties that it may emerge during the course of the Hearing that there was prejudice to the Respondent that had not been identified in allowing the amendment but that could be relevant to our decision whether to extend time and secondly, prejudice is not the only factor in deciding this time limit point. We would also need to take into account the length of and reasons for delay, giving the Claimant the opportunity to give evidence on those points.

Hearing

20. Witness statements were prepared by the Claimant, Mr R Kockelbergh (consultant urological surgeon and honorary professor of urology based in Leicester), Mr J Crew (consultant urologist in Oxford), Mr M Woodward (consultant paediatric urologist in Bristol), Ms Z Gall (consultant urologist in Stockport), Mr R Mills (urological surgeon), Mr C Shukla (consultant urological surgeon), Mr D Hodgson (consultant urologist in Portsmouth), Mr D Thomas (consultant urological surgeon in Newcastle-upon-Tyne), Mr W Ryan (Consultant Orthopaedic Surgeon in Bolton and Chair of the Respondent's Internal Quality Assurance committee ("IQA") from 2020, his predecessor being sadly deceased) and Mr A Dickinson (consultant urologist in Plymouth and Chairman of the Urology Intercollegiate Specialty Board ("ISB") from 2019 to 2022). We heard oral evidence from all of these witnesses, except Mr Thomas and Mr Woodward who were unavailable – see further below. We therefore attached less weight to their evidence as it was not tested. Some of those witnesses who gave oral evidence did so remotely, without objection from the Claimant.

21. The parties had agreed a bundle of documents and the Claimant had also compiled a supplementary bundle – the Respondent had no objection to that being before us. Together the bundles totalled around 1,000 pages. We read around the first 100 pages, being the pleadings and related documents. Counsel suggested some pre-reading of a limited number of further documents. Having reviewed them, we made clear that we had only skim-read them before hearing evidence, as it did not seem to us necessary to read them in full. We read the statements, totalling 95 pages, but not all of the documents referred to in them, for the sake of time. We made clear repeatedly that it was for the parties to take us to any documents they wanted us to consider in detail. One additional document was handed up by the Respondent during the Hearing, namely a set of examination questions from before 2019. It was added to the bundle without objection.

22. The Hearing progressed as follows:

22.1. Day 1 was given over to agreeing the issues and to the Tribunal's pre-reading.

22.2. The next seven days were devoted to hearing evidence (plus the amendment application). It was not possible for the Tribunal to sit on 5 June 2023 due to my having another pre-arranged judicial commitment.

22.3. The parties did not attend on Friday 14 July, which gave both counsel an opportunity to prepare written submissions after the evidence had been completed. The Tribunal used that day to review its notes of evidence and to start (but not complete) its findings of fact.

22.4. We heard oral submissions on Monday 17 July. Both counsel provided detailed written submissions which we read beforehand.

22.5. The Tribunal's deliberations took place on 18 and 19 July, including completion and review of our findings of fact.

22.6. Two further days, 20 and 21 July, were given over to writing up the Judgment and Reasons.

Facts

23. Our findings of fact were based on the above material. References to page numbers below are to the agreed bundle, any such references pre-fixed with "SB" being a reference to the supplementary bundle. Alpha-numeric references are to witness statements, for example AG5 means paragraph 5 of the Claimant's statement and AD4 paragraph 4 of Mr Dickinson's statement. As with any Tribunal decision, we do not detail in our findings all of the matters rehearsed in the evidence. Instead, we have focused on those which seemed to us of most relevance to the issues we have to determine.

Background

24. The Respondent is an unincorporated body which acts on behalf of the four surgical Royal Colleges of Edinburgh, Glasgow, England and Ireland. The four Royal Colleges work together on matters relating to standards and assessment through a number of intercollegiate committees. The Respondent is one of these committees. It is responsible, in line with General Medical Council ("GMC") statutory requirements, to the four Royal Colleges for the supervision of standards, policies, regulations and professional conduct in respect of the Intercollegiate Specialty Board ("ISB") examinations, one for each of the surgical specialties, including urology. Each ISB is responsible to the Respondent for ensuring specialty examinations reflect the curriculum being undertaken by surgical trainees in the United Kingdom and Ireland.

25. The Claimant is a urologist of Egyptian nationality and Coptic ethnicity. His first language is Arabic. He graduated in Egypt in 1999, started training in urology in 2002, from 2005 to 2008 worked as a registrar in the International Hospital for Urology and Nephrology in Cairo, became Senior and Chief Registrar until 2010 and was then a specialist urologist. In this last role in Egypt, he ran

clinics, ward care, operating theatres and worked in all subspecialities of urology. In his view (AG6-7), his overseas training gives him a broader spectrum of experience than people who train in the UK. He says for example that UK-trainees would not typically be able to do open (as opposed to keyhole) surgery. From 2010, he also had his own private urology practice.

26. In 2013 he passed an examination to become a Member of the Royal College of Surgeons (MRCS), England. He worked in Dudley from 2014 as a Urology Senior Clinical Fellow and became a Urology Specialty Doctor in Birmingham in 2015. From 2016 onwards, he was approved by the clinical director, after a multi-consultant assessment, to operate and run clinics and diagnostic sessions without supervision.

FRCS-UROL Examination

27. There are two routes to being entered on the Specialist Register held by the GMC so as to be entitled to be appointed and work as a consultant urologist in the NHS. Candidates employed by the NHS in surgical speciality training posts (after two years' foundation training and two years' core training) must complete higher surgical training, run by the Joint Committee on Surgical Training ("JCST"), at which point – after five or six years, depending on the specialty – they receive what is known as the Certificate of Completion of Training ("CCT"). They do not have to be UK nationals to enrol on to the JCST training. Once they enrol, they are given a National Training Number. According to Mr Dickinson, they are essentially supernumerary, attending regular training, conferences and benefiting from peer support. The Intercollegiate Surgical Curriculum Programme ("ISCP") for Urology at page 654ff shows the process a surgical trainee will go through, which includes six work-based assessments. In order to be entered on the Specialist Register, a trainee must as part of the JCST training pass professional examinations including the FRCS examination, as well as being assessed in other ways such as by observation of their surgical work. As Mr Ryan explained, some senior doctors from overseas are permitted to complete the JCST training in a shorter period, because of their experience, though it requires a minimum of 18 months.

28. The alternative route to being entered on the Specialist Register, that is for those not in training posts, is to obtain the Certificate of Eligibility for Specialist Registration ("CESR"). The CESR does not necessarily require the person to pass the FRCS exam, though it can be – and, as Mr Ryan confirmed, most often is – taken by individuals in this position. If it is not, then they will have to show convincing evidence that they have demonstrated substantially equivalent knowledge; having almost passed the FRCS exam would not count. At page 666ff, there is a GMC document which summarises the two routes and the steps required for registration via each. At page 674 it is stated that a CESR candidate must demonstrate knowledge to the standard of the CCT curriculum, otherwise supporting evidence "must be very strong indeed". There are some UK nationals who take up non-training posts and therefore have to pursue the CESR route to get on to the Specialist Register, but Mr Ryan told us that most UK nationals are successful in securing training posts. The CESR route requires provision of three references. The Claimant's references are at pages SB1 to SB10. To our reading, they commend the Claimant very highly. One was from a clinical director and two from supervising consultants.

29. It is for the GMC to grant or deny an individual's application for entry on to its Specialist Register. The Claimant says (AG60) that it is "practically impossible" to be successful in such an application without the FRCS exam. He did not apply for an NHS training post for various personal reasons, including financial considerations and the fact that he did not for family reasons wish to work in various locations. Mr Ryan told us that of around six CESR applications per year in Urology on average, just over 50% obtain it.

30. At page 251ff is a document in which the GMC sets out the standards for postgraduate curricula. It emphasises the importance of patient safety as the first priority and stresses the need to maintain standards across the UK, to encourage excellence, and to embed fairness. The Intercollegiate Surgical Curriculum in force in 2019 can be seen from page 221ff, mirroring the GMC's standards. It is not set by the Respondent. The Claimant's view is that the Respondent did not properly translate these principles into the examinations which he undertook, though he has no issue with the NHS trainee curriculum being the measure against which he was examined. His issue is that how he was marked was unfair, that is, not a fair reflection of his performance in the examination process.

31. The Respondent's Internal Quality Assurance Committee ("IQA") is responsible for ensuring the FRCS examinations are reliable and delivered in the appropriate way. The examination format can only be changed with the permission of the GMC. Section 1 of the FRCS examination is a computer-marked multiple-choice assessment. The Claimant passed this first time in January 2018, scoring 76% against a pass mark of 66%. There is no human judgment in the assessment of this section. The pass mark varies each year. The Claimant asserts that the Respondent varies the pass mark so that more UK nationals pass than would otherwise be the case. Mr Dickinson told us that the pass mark is set according to an assessment of how difficult it has been, using a statistical analysis called the Angoff technique. We prefer Mr Dickinson's considered evidence to the Claimant's speculation. The success rate for trainees and non-trainees in section 1 over the period 2016 to 2019 (page 591) is 73.9% and 49.5% respectively. This Claim is entirely concerned with section 2 of the FRCS-UROL (that is, urology) examination, which will be referred to in the rest of these Reasons as "the Examination".

The Examination process

32. The Examination is a scenario-based oral test, which candidates have a maximum of four attempts to pass. It is intended to be a test of judgment and ability to solve problems (RK2) and is therefore not susceptible to binary answers or assessment. Ms Gall says (ZG3) that it seeks to assess whether candidates can independently and safely manage problems across all different aspects of urology, whilst Mr Shukla says (CS5) that candidates may have relevant knowledge, but the Examination is about applying it to clinical scenarios, which requires judgment and experience. It thus represents real clinical practice (DT24). The Claimant accepts that marking such scenarios involves an element of subjective judgment, such that two examiners could mark the same answer differently, and that this is a common way of assessing those in medical practice. It is agreed that the Examination is not easy to pass.

33. Each examiner must have had several years practising as a consultant and there are certain criteria that have to be fulfilled to be appointed – page 301. We

have not considered the criteria in detail, but note that it is said examiners must "be able to demonstrate a policy of courtesy, fairness and non-discrimination". They also undertake some induction training, including in relation to equality and diversity – see page 383ff. The training includes reference to unconscious bias, which the presentation slide at page 389 suggests can be overcome by being aware of it, by not rushing decisions and by justifying decisions by evidence. Ms Gall told us that the diversity training encourages and instructs examiners to focus on the content of a candidate's answers. If a candidate's first language is not English and there is any difficulty understanding questions, examiners are trained to make adjustments such as breaking down the question into parts, checking understanding and simplifying language. Mr Ryan has never seen any report about an examination (as to which, see below) raise concerns about a candidate's ability to communicate in English, though Mr Dickinson told us that some candidates have raised concerns about difficulty in understanding certain examiners. That was not the case for the Claimant.

34. There are eight parts to the Examination, known as "stations", each assessed by two paired examiners (one usually a more experienced examiner than the other), over 20 minutes each, looking at two ten-minute scenarios. Each examination period is known as a "diet". Typically, there will be two or three sets of examinations going on at any one day, each testing around 16 candidates, so up to around 48 candidates in total.

35. Different examiners are deployed to each station, by the ISB Chair – Mr Dickinson in this case. Because examiners may have specialist subjects, he would try to rotate them to create some variety, changing the pairs for each of the two days of each diet (AD13). A candidate should not be examined by the same examiner twice in the same diet (AD32). One or two assessors, appointed by the IQA, observe each examiner during each diet and can be quite critical (JC6) when assessing examiner performance by way of their subsequent reports, for example saying that they should have moved through a scenario more quickly. Mr Mills says (RM27) that they also assess whether examiners are being attentive to candidates.

36. One of the examiners allocated to each station asks the questions, whilst the co-examiner observes. As Mr Kockelbergh explained, the examiners briefly discuss what was said by the candidate once the 20-minute station is concluded, essentially for the benefit of the lead examiner who will not have taken the same level of notes because s/he will have been focused on interacting with the candidate. This brief discussion is to ensure nothing significant that the candidate has said is missed though there is no discussion of the marks that are to be given. Both examiners then score the candidate's performance separately, on marking sheets. The Claimant's marking sheets for November 2019 can be seen from page 714 onwards. Any difference of more than one mark between examiners has to be considered further by the Respondent, and examiners must then note their reasons. There is around 10 minutes after the end of each station in which to allocate a mark and get ready for the next candidate to arrive.

37. Candidates are assessed on the basis of them being a day 1 consultant in the UK. We will return to this below. It does not mean (RK3) they have to know everything, but if making decisions, they should be safe. Mr Crew says (JC3) that it is about whether a person can practice independently, know their limitations and have a basic understanding of urological emergencies, whilst Mr

Dickinson says (AD24) that a consultant is day 1 ready if they can manage most patients, know how to manage their juniors, and also be able to ask for senior help. They do not need to be good at everything, but competent, safe and of the standard of knowledge set.

38. A group of consultant surgeons prepare a bank of questions for each Examination diet beforehand (the questions will have been mapped against the ISCP, and those who have drafted them report on this process to the ISB board). At each diet, on the day before examiners start seeing candidates, each subject lead examiner selects questions from the bank and there is then a calibration (standard-setting) session at which expected responses are reviewed by each examiner who will be assessing candidates in the particular subject area. This discussion takes several hours and if someone does not attend it, they cannot examine (DT9). The questions are discussed so as to set the pass and fail points for each, including what would be a fair pass, a good pass and an excellent pass. Each question thus has a template or marking descriptor of expectations for the answers, which is then shared at the start of the exam day with each examiner who will be marking in that subject. The starting questions set the scene and the follow up questions are more complex (WR13). More than one of the Respondent's witnesses informed us that the marking descriptors for the questions are left with the Respondent's administrators after each diet. No such documents, whether from November 2019 or otherwise, were available to the Tribunal. The Respondent could not explain why.

39. For the November 2019 diet, Mr Dickinson introduced a handwritten grid, as a pilot, to ensure that whilst at the station examiners had some way of referencing what answers related to which grades (from the standard setting exercise), rather than that being written on the question sheets as had previously been the case. No such grid was available to the Tribunal either; Mr Dickinson does not think they will have been retained. Any amendments to the grid during the calibration discussion were handwritten on to it by each examiner. Question sheets for current examinations provide more clarification of which part of the marking sheet relates to which part of the question.

40. When a candidate arrives at the first station at which they are to be assessed, they are welcomed by the co-examiner, who checks their candidate number and that they are at the right station, explains the Examination process and then hands over to the lead examiner for that station. At each of the eight stations, as can be seen from the marking sheets, a candidate is assessed in three areas, although a number of the Respondent's witnesses say that the distinction is somewhat artificial. The three areas (generally referred to in this Hearing as "domains") are:

40.1. Knowledge/Comprehension/Organisation ("Knowledge").

This is about background knowledge of urological conditions. Mr Shukla says (CS21) that it is about being able to identify what the patient's problem is or could be, and the ability to organise one's thoughts.

40.2. Application and analysis ("Application").

This is concerned with how knowledge is applied to the specific situation, that is patient management on the ward. Ms Gall says (ZG16) that it is putting

knowledge into practice to understand how it fits with what a patient describes, whilst Mr Shukla says (CS21) it is forming a diagnosis and plan.

40.3. Synthesis and judgment ("Synthesis").

Each witness describes this domain slightly differently, which Mr Mills accepts means they may mark it in different ways. Mr Crew says that it is about making safe and good decisions (JC7 – a complication or something unexpected); Ms Gall says (ZG16) that it is about creating a sensible management plan (she told us it is about bringing general knowledge and knowledge of the patient together for this purpose); and Mr Shukla says (CS21) that the focus is on what treatment option the candidate would recommend. Mr Mills told us it looks at "deeper knowledge" (RM24), namely whether the application of knowledge – with which it very much overlaps – is undertaken safely. This domain has now been replaced by one headed, "Professionalism".

41. An average of 6 marks in each domain is required in order to pass (the highest mark being 8 and the lowest 4). Mr Shukla told us that it is difficult to get a 6 consistently, whilst Mr Mills said that a 4 is very unusual and 8 uncommon. This results in a pass mark of 576, from a combination of eight stations, two examiners at each, both marking in three domains, for each of two scenarios, with a pass mark of six. Accordingly, a candidate could fail several stations and still pass overall and end up practising in the field in which they failed. A mark of 576 is said to equate to 75%, the lowest mark is 384 (50%) and the highest 768 (100%). Mr Ryan was not sure why the Examination is marked in this way (rather than, say, 1 to 8) but believes there is research to suggest that it encourages examiners to use a breadth of marks. He added that the critical aim of protecting patient safety is built into the marking scheme by a mark of 6 equating to what the question-setters think constitutes safe practice. In the remainder of these Reasons, we record each set of three marks simply using three numbers, for example 566 refers to a mark of 5 for Knowledge, 6 for Application and 6 for Synthesis.

42. The three domains on each marking sheet do not match the criteria for assessment on the generic marking descriptors at page 216 (dated April 2016). In respect of prompting of candidates, they say requiring frequent prompting leads to a score of 5, minimal prompting is a 6 and "fluent responses without prompting" is a 7. In respect of quoting literature, they say "providing supporting evidence and familiar with literature" is an indication of a 7, whilst "had an understanding of the breadth and depth of the topic and quoted from literature" indicates an 8.

43. Mr Crew says (JC15) that provided most factual questions are answered correctly, candidates will probably get a 6. Someone who is struggling or needs prompting to move the test along will get a 5. He says at JC19, "you think of the scores as you go along" and if a candidate is a 6 or 7, "you get that impression quite quickly". He explained in oral evidence that this is a reference to early, not first, impressions, meaning that if a candidate answers well early on in a scenario, the examiner knows they can push them on to more difficult questions because they will be at least a 6. Several of the Respondent's witnesses stated that examiners want to pass every candidate. Ms Gall told us that most candidates who fail do so because they do not involve the patient, the patient's family and the candidate's colleagues in decision-making.

44. The Respondent's briefing document on marking the Examination is at pages 219 to 220. It says that marks should be justified "with explicit reference to the marking descriptors where possible (especially if you assign a 4 or 5)". It goes on to say, "Please try to justify your marks from the marking descriptors … this justification of marks … should give examiners confidence in the marks they have assigned". Ms Gall agrees that it is not good practice not to write anything on the marking sheet, especially where the score is below 6, but notes that it is difficult for the examiner who is asking the questions to take any notes at all. Mr Ryan says that "ideally" a justification for a mark would be recorded on a marking sheet but added that examiners mark immediately after the viva and can recall what they have heard. Mr Crew told us that as a result of that immediate marking, even if a note on the exam sheet is incorrect, that would not influence the mark a candidate is given.

45. Mr Ryan also accepted that there should be comments in most, if not all, of the three domains for each scenario, certainly where a 4 or 5 is given. A descriptor should be referred to, he and Mr Dickinson said, "where possible", as per the briefing document, Mr Dickinson saying that some examiners feel it unnecessary to justify a pass, so that reference to descriptors is less relevant for a mark of 6, 7 or 8. Mr Shukla told us that not all examiners approach the marking in the same way, for example he is more lenient than some on whether someone has been prompted, but this is something which is open to interpretation, for example the significance of asking a candidate, "Anything else?", which he would not take as prompting but which some would. In his view the 96 sets of marks a candidate gets overall (6 marks from each of two examiners at eight stations) evens things out.

46. Before each diet, as well as the calibration of the questions described above, there is a board chair's briefing for examiners (WR15), a briefing for the candidates (WR16) and a training session on equality and diversity for the examiners, delivered by one of their peers, in either case chosen by the ISB chair. The training for November 2019 appears to be that in the PowerPoint from pages 428 to 457 (we were only taken to a small number of those pages). It was suggested by one witness that Mr Shukla delivered it, but he himself could not be sure, though he has delivered the training on two occasions. He pointed out to us page 443 which includes photos of different people and told us that the training discusses unconscious bias based on appearance and the importance of not jumping to conclusions. He accepted, similarly to Mr Crew, that he gets a rough impression of candidate performance in the first couple of minutes but says that examiners are trained to keep an open mind.

47. At page 432, the training presentation says, "You all know that despite extensive mandatory training programmes there is still differential attainment" including in exam results. Mr Ryan was unsure what this refers to but told us that the focus is on the differential attainment of those in training posts and those who are not. The Respondent accepts that candidates in training posts perform significantly better in the Examination than those in non-training posts, something we will return to below. It says that the fact that the Claimant was not in a training post was a material difference between his circumstances and the circumstances of those who were.

Claimant's initial Examinations

48. The Claimant attended courses (AG18) in preparation for the Examination and studied and practiced for it with four colleagues. A Mr K Murtagh and A Ms S La Touche were UK nationals and surgical trainees and passed it first time. Mr S Nafie and Mr A Christaphides were non-UK nationals (though both British citizens with dual nationality and surgical trainees) and passed on the second and third attempts respectively.

49. The Claimant failed his first attempt at the Examination in May 2018, by sixteen marks (scoring 73.4%). A surgical trainee of Syrian nationality, Mr A Zrek, also failed the first attempt. He, and Mr Nafie passed on the second attempt on 15 November 2018, an Examination which the Claimant failed by five marks (scoring 74.4%). The Claimant says that during this second Examination he raised with Mr N Burgess, then Chair of the ISB, some concerns about two of the examiners, Mr Ravi and Mr Kockelbergh, saying that they intentionally slowed him down so that he did not reach the end of the clinical scenario, repeatedly challenged his answers, and frowned at him. The Claimant says that Mr Burgess advised him to complete the Examination and to write to him if the result was not as expected. The Claimant says he did so but got no reply and that Mr Burgess now denies that the conversation took place. We did not hear from Mr Burgess nor were we directed to any documentation or correspondence produced by him and thus there was no challenge to the Claimant's evidence on this matter. The Claimant's broader point was that the Respondent's evidence cannot be trusted.

50. The Claimant also raised concerns about Mr Kockelbergh in his feedback form after the second Examination, saying that he denied the Claimant had answered certain questions, frowned at him, and failed to shake his hand at the end unlike the other examiner. Mr Kockelbergh was not spoken to about the Claimant's concerns, which he thinks now a little surprising. He says that before the Covid-19 pandemic it was commonplace to shake hands when greeting candidates, though with the passage of time he cannot recall whether he did that with the Claimant. Examiners are however positively discouraged from giving positive or negative indications after the exam, including shaking hands. He cannot recall frowning and makes no apology for being challenging, which he says is what the Examination is intended to be.

51. The marks given by Mr Kockelbergh and his co-examiner at the second Examination can be seen at pages SB23 and SB25. Mr Kockelbergh gave him 656 for scenario 1 and 666 for scenario 2. His unnamed co-examiner gave him 666 and 777. Mr Kockelbergh's mark sheet (page SB25) leads him to believe that the Claimant got the initial question wrong so that it is likely he asked him about it several times to ensure he was saying what he meant to say. "ABCDE", noted on the mark sheet, refers to a primary survey for an acutely ill or injured patient. Mr Kockelbergh put a cross next to it, which he says supports his scores. He finds his colleague's mark surprising. Mr Kockelbergh wrote no comments regarding scenario 2 on this occasion, giving a mark of 666, whilst his colleague scored it 777. Mr Kockelbergh told us that when he is the lead examiner, he only writes a note where there is a problem with an answer or something very good, so that if he writes nothing that almost always signifies straight sixes.

52. The Claimant sat the Examination for the third time in May 2019 and failed by three marks (scoring 74.6%). After each Examination, he had informal reviews of his performance with his principal referee.

Claimant's final Examination, November 2019

53. The Claimant failed his fourth and final attempt at the Examination by two marks (scoring 74.74%) in November 2019. It is that outcome which is the sole focus of his complaints in this Claim, though he says certain examiners who had marked him unfairly before, including Mr Kockelbergh, were part of this assessment and should not have been. He says (AG32) that had different examiners been provided, he would have been successful.

54. We recount first the examiners' evidence about the Examination and then the Claimant's challenges to their marking. None of the examiners recall the occasion, all saying that they only recall the exceptionally good or notably poor candidates. They were thus relying on the mark sheets when preparing their statements and giving oral evidence. They each had several years of examining experience by 2019.

55. The scores given to the Claimant were as set out in the table below. Some of the examiner names were only made known to us in Mr Briggs' written submissions. As a result, we do not know for stations 4, 6, 7 and 8 which of the examiners gave which scores. We were not taken to scores for any other candidate, whether given by the same examiners as marked the Claimant or otherwise.

Station	**Examiners**	**1st examiner scores**	**2nd examiner scores**	**Difference in examiner scores**	**Page numbers**
1	R Kockelbergh and J Crew	666/667	665/666	2	714/716
2	R Mills and C Shukla	566/ 666	676/666	2	718/720
3	Z Gall and M Woodward	655/655	565/665	1	722/724
4	D Summerton and A Mathur	666/677	777/777	4	726/728
5	D Hodgson and D Thomas	666/676	666/666	1	730/732
6	R Ravi and R Webber	666/666	666/666	0	734/736
7	T Sami and K Walsh	665/655	666/655	1	738/740
8	R Napier-Hemy and R Morley	667/666	666/666	1	742/744

56. It can be seen therefore that there was an overall difference of 12 marks between the two sets of examiners. This is out of a total, if taking what is needed to pass, for all eight examiners and all eight co-examiners respectively, of 288.

R Kockelbergh

57. Mr Kockelbergh was the co-examiner for Station 1. He did not recall the Claimant from the previous year. His evidence in relation to the first scenario, which concerned radiation cystitis, was as follows:

57.1. He gave a mark of 6 for Knowledge, saying at RK24 that the Claimant "looked reasonably sensible". He goes on to say that the Claimant "did not do well with histopathology and did not describe healing problems but had a reasonable level of knowledge", though he did not really know other areas where Mr Kockelbergh and Mr Crew tried to probe him, such that Mr Kockelbergh concluded, "Had he had no significant deficiencies in knowledge, he would have scored a 7". In his oral evidence Mr Kockelbergh told us that the Claimant got all the "easy stuff" correct in this scenario but did not know the histopathology of radiation cystitis and some of the treatments the Claimant referred to were historical. He correctly identified embolization as a treatment and that the patient had a bladder tumour. Had he described the risks and benefits of various treatments, and when each is appropriate, he would have got a higher mark.

57.2. In relation to the mark of 6 for Application, Mr Kockelbergh says at RK26, "he had to be prompted to talk about embolization and did not know about hyperbaric oxygen". He now accepts that the Claimant did <u>not</u> have to be prompted about embolization and <u>did</u> know about hyperbaric oxygen, though he had to be prompted about it. Mr Kockelbergh wrote on his mark sheet, "dilated ureter, prompted", which he thinks likely means the Claimant was asked what it was on a scan. At RK26 he describes the Claimant as, "Nowhere near a good pass but acceptable". He stands by that, saying that the Claimant only got through half of the question and that "slick candidates" move through the scenario quickly.

57.3. As for the mark of 6 for Synthesis, in his statement he says, "The Claimant talked about resection and wanting to do a ureteroscopy, which I thought was incorrect". This is reflected in his notes, where he has written "URS" and put a cross next to it. He told us that a ureteroscopy was pointless and an almost certainly impossible treatment in this scenario, that it was not required and would almost certainly fail. He says at RK29, "He did not get into any discussion about bladder cancer which is what nearly half the question was about".

58. Mr Kockelbergh's evidence in relation to the second scenario was as follows:

58.1. He gave the Claimant a 6 for Knowledge, saying at RK30 that he showed "some indecision about endoscopy being inconsistent" and was "not fluent in the expression of his knowledge and he has made some inconsistent decisions about endoscopy". He says he would have been given a higher mark if he had been more decisive and consistent about his plan for the patient.

58.2. In relation to the 6 for Application, Mr Kockelbergh says that this was a difficult question (RK32). He says that the Claimant struggled "with the consent process" and was indecisive on whether to wake the patient. He describes the Claimant as good enough, but not exceptional.

58.3. As for the 7 for Synthesis, he says that the Claimant quoted more of the literature on the comparison of different types of surgery.

59. Overall he thus regarded the Claimant as a pass candidate. In oral evidence, he accepted that the Claimant identified an inconsistency between the histology and an endoscopic image. He accepts that this is what "inconsistency" refers to in his notes and that accordingly his references to inconsistency in his statement (RK30 and RK32) regarding scenario 2 are incorrect. He could not recall in oral evidence what the Claimant's indecision related to. He thinks he gave him a 7 for Synthesis because he probably described a clinical trial in some detail.

60. Mr Kockelbergh says he has no knowledge of the nationality or ethnic background of any Examination candidate, although their names may give some idea.

J Crew

61. Mr Crew was the lead examiner at station 1. He wrote only one comment on his mark sheet at page 716, telling us that time pressure means he does not write anything where the mark is a 6. For scenario 1 his evidence was as follows:

61.1. He gave a mark of 6 for Knowledge, as the Claimant talked about patient history and investigation, the need for a biopsy and catheterisation.

61.2. He gave the Claimant a 6 for Application, as he knew the basics of more complex management issues with radiation cystitis. What Mr Crew wrote on his mark sheet was in fact that the Claimant did <u>not</u> know any management for radiation cystitis, in contrast to Mr Kockelbergh who wrote down several possible treatments the Claimant had mentioned. Mr Crew accepted in oral evidence that there is a plain contradiction both between his evidence and the mark sheet and within paragraphs 23 to 25 of his statement in which he states both that the Claimant knew about management for radiation cystitis and that he did not. He agrees that both his statement and what he wrote on the marking sheet were inaccurate and that in hindsight he would write different things to justify his marks. The explanation for his mark he settled on was that the Claimant knew what radiation cystitis is but there were deficiencies in his knowledge of how to manage it.

61.3. In respect of the 5 for Synthesis, Mr Crew said in his statement that he could not remember why he marked the Claimant down but that Mr Kockelbergh's notes state that he needed prompting. He would have achieved a higher mark if he had more fluidity in his answers. In oral evidence, Mr Crew said that the Claimant was given a 5 because of his understanding of radiation cystitis and its management, also referring to Mr Kockelbergh's notes about histopathology and healing problems. It also seems that when writing his statement (JC25) Mr Crew misunderstood Mr Kockelbergh's mark sheet note about a ureteroscopy, Mr Crew saying that the Claimant failed to mention it, when Mr Kockelbergh's point was that it was wrong to mention it.

62. As to the second scenario, Mr Crew said at JC27 "the lack of writing on my notes implies this candidate was absolutely safe".

63. Mr Crew says at JC31-2 that he has no knowledge of candidates' race but can "obviously see this from their appearance if they are Asian or African or if they have an accent". He did not have any information about a candidate's training or whether they are in a training or non-training post. The Claimant does not accept that. His case is that his name, colour and accent make clear that he is a non-UK national. He also mentioned a question put to him by one (unidentified to us) examiner at one of the Examinations about where the Claimant had carried out a particular procedure. The Claimant told him it was abroad. The Claimant also says that his own results and the disproportionately lower pass rate for non-UK nationals shows that examiners know whether someone is in a training post or not.

R Mills

64. The first scenario at Mr Mills' station, for which he was the co-examiner, was concerned with prostate cancer. His evidence about his marks was that on the mark sheet he had heavily underlined the words, "some prompting". He says at RM33, "A candidate generally should not require prompting [about] history taking and examination for prostate cancer as it is a common clinical presentation".

65. The second scenario was concerned with penile cancer, which is much rarer. Mr Mills gave the Claimant 6 for Knowledge, saying that the Claimant identified the need to look at risk factors, examination and a biopsy, describing the Claimant as demonstrating sufficient knowledge to pass.

66. Mr Mills told us that the marking descriptors are useful in deciding what mark to give, but as long as a mark is justified in some way allied to them, that should suffice. At RM38 he says that the Claimant's marks were "not great across the board" but that he fell down particularly in paediatrics and bladder dysfunction. He describes as "impossible" the idea that he could work out from the Claimant's colour, accent and name that he was an overseas doctor.

C Shukla

67. Mr Shukla was the lead examiner for station 1. He says at CS27 that for this station, concerned with prostate, testis and penis urological oncology, candidates needed to be able to assess, investigate and have reasonable knowledge of their management options and then refer the patient to experts. Mr Shukla's evidence in relation to his marks for scenario 1 was as follows:

67.1. Whilst he gave a higher mark for Knowledge than Mr Mills, he noted that Mr Mills is an expert in prostate cancer. He told us he was speculating that this was the reason for the lower score but imagines that Mr Mills would want a candidate to deliver all of the basics without prompting. He accepts that had the Claimant required significant prompting he would have noticed. As to his own mark, he told us that a mark of 7 would require identifying certain things on the scans and making referrals, perhaps even backed up by the literature.

67.2. He thinks it was the fluidity (by which he means confidence and fluency) of the Claimant's answer which led to him giving a 7 for Application, not an 8. As for Synthesis, Mr Shukla told us that knowing the name of a research paper is only a 6, showing a basic understanding of it is a 7, and showing a detailed understanding of it is an 8. He does not recall – as the Claimant asserts –

whether the Claimant mentioned a particular trial, even though he accepts that Mr Mills' notes suggest he did. In Mr Shukla's view, it is mentioning the benefits of the trial that would lead to a 7.

68. The second scenario was Mr Shukla's field of interest; he scored the Claimant 666. He says at CS35 that if a candidate moved through it fluently, for example measuring phallic length, or quoting certain literature on penile preservation data and outcomes, that would score a 7. If they said it looked like penile cancer and eventually got to the basics such as risk factors and management, that would be a 6. The Claimant was in that latter category.

69. Mr Shukla says at CS40 that a candidate's ethnic background is sometimes obvious when you see them but their nationality and whether they trained in the UK is not evident from that information.

M Woodward

70. Mr Woodward was the lead examiner at this station. He says at MW19 that most candidates will never do paediatric urology, so that the pass threshold for this station was relatively low. Scenario 1 was concerned with a boy with an undescended testicle. Mr Woodward's written evidence in relation to his scores as follows:

70.1. The Claimant scored 5 for Knowledge because (MW32) he was confused about the fertility risk which means he would not have been able to give accurate information to the parents. Noting that Ms Gall gave a 6 in the same category (see below), Mr Woodward suggests that there would have been some aspect she was impressed with which did not strike him, or it could have been her lesser paediatric knowledge which led to the higher score.

70.2. In respect of the 5 for Synthesis, Mr Woodward says that the Claimant suggested a laparoscopy, which was a significant error – it would involve doing an operation the candidate was not trained for, would not have appropriate equipment for and did not have parental consent for.

71. Scenario 2 was a child with night-time wetting. Mr Woodward commented on his marking sheet that essential points were mentioned. In respect of the mark of 5 for Synthesis, he says that candidates are expected to start with simple tests and work up from there whereas the Claimant suggested proceeding straight to urodynamics. Mr Woodward describes this at MW41 as a significant error.

72. According to MW45, two of Mr Woodward's consultant colleagues, whom he appointed, graduated from the biggest university in Cairo. One or two fellows/registrars per year are taken from there also. In his view (MW46), clinical experience is often higher in overseas trainees, though Mr Shukla told us that it is not a given that someone's surgical experience reflects their competency. Mr Woodward also says at MW47 that the pass rate is higher for those in UK training posts because they have better exam technique. Mr Ryan accepts that as a fair point to make, given the supervision trainees receive, including with regard to passing the exam.

Z Gall

73. Ms Gall was Mr Woodward's co-examiner. She has not practised paediatric urology although it was part of her standard training and she did some focused training in it in 2010/2011 when she was considering it as a sub-specialty. Mr Woodward was thus the specialist for this station. Ms Gall accepts that in such a situation, the non-specialist may need confirmation that they have got something right, or indeed that the candidate got something wrong. She points out however that the bedwetting scenario in particular was not complex.

74. In respect of the first scenario, she says that the Claimant was confused about the fertility risk. She accepts that why he did not meet standard is not fully set out in her notes on the mark sheet, though she says at ZG23 "I probably meant he didn't get it quite right or didn't give us confidence that he knew that scenario well enough". She comments further:

74.1. She scored him a 6 in the first domain as he had some knowledge but did not give confidence that he could counsel the parents appropriately. A higher mark would require explaining the risks and benefits of various treatments.

74.2. He scored 5 for Application because he mentioned atrophy only after prompting.

74.3. He scored 5 for Synthesis – she would have marked this higher if he had given more confidence that he could counsel the parents appropriately.

75. As to the second scenario, Ms Gall says that the Claimant had some understanding of different types and causes of bed-wetting, but it was a major error to suggest urodynamics studies before simple things like keeping a chart of how often the child uses the toilet. She added:

75.1. The Claimant scored 6 for Knowledge as his book knowledge was strong.

75.2. He scored 5 for Application as he knew a lot but did not come to the right conclusion about how to investigate the patient.

75.3. He scored 5 for Synthesis as the manner of the investigation he identified made it appear he did not have experience.

76. Ms Gall was the only witness whose marking analysis we were taken to (see page 472ff). This shows that she gave more than the average number of 5's and 7s, and less than the average number of 6s. Of the 5s, she gave 14 to candidates whose first language was not English, and 7 to those whose first language was English, though she also gave more 6s and 7s to Asian candidates than to White candidates.

D Hodgson

77. Mr Hodgson was the co-examiner for the stones/UTI station. He gave the Claimant 6 for Knowledge in scenario 1 and says he may have scored higher if he had shown fluidity of thought, challenging Mr Hodgson as an examiner, by for example introducing supporting literature. He may have scored a higher mark than 6 for Application by demonstration of fluidity of thought. As for scenario 2,

the Claimant scored 7 for Application, Mr Hodgson noting that there was a discussion of different treatment options.

78. Mr Hodgson told us that whilst it is important to justify marks from the descriptors, choosing a mark to assign for a question pre-supposes the examiner has had reference to the descriptors without having to write that down each time.

D Thomas

79. Mr Thomas was the lead examiner for the stones/UTI station. He gave 6s in all three domains for both scenarios, saying at DT35 that it is difficult to say specifically why. He says that the Claimant did not excel but was a safe pass, and that (DT36) he would probably look for a candidate to move through the basic questions quickly and get into some more complex areas or quote literature in order to score more highly. He says at DT38 that where there is an average candidate, he writes very little on his notes. He wrote "safe/stent" for Application on scenario 2. Mr Hodgson told us that "safe" and "competent" are similar.

Claimant's challenge to his scores

80. The outcome of the Claimant's fourth Examination was given to him in a letter of 26 November 2019 from Mr Burgess (page 746). The Claimant says after receiving it he wrote to the Respondent's President and to the Chair of the ISB (Mr Dickinson) without any reply. Mr Dickinson does not recall any such communication. It seems likely on this basis that the Claimant did send such communications, though we were not taken to them. The Claimant's case is that he has identified a number of errors and inadequate explanations in the marking of the Examination, all of which he raised in his formal appeal to the Respondent which was sent on 4 December 2019, with further observations following on 19 and 20 December 2019 – we were only taken to short extracts from this material. We have already set out some of the Respondent's comments on the relevant assessments above. We now set out the Claimant's challenges to the scores he was given and some additional points in relation to the same raised by the Respondent's witnesses.

Station 1 (Mr Kockelbergh and Mr Crew)

81. The Claimant says that Mr Kockelbergh's mark of 6 for Knowledge on scenario 1 should have been a 7, given Mr Kockelbergh's notes on his mark sheet setting out supporting evidence. Mr Kockelbergh told us that he writes a record of the whole viva so that what is written in a particular box on the mark sheet may not relate to that domain, because a single answer may cover more than one. Mr Crew had a similar view. For what he described to us as a "fairly standard candidate", he relies on what he has written on the mark sheet. He also emphasised to us that there is only a short time to mark.

82. The Claimant points out that for Application on scenario 1, Mr Kockelbergh referred to the Claimant's indication of various lines of treatment of radiation cystitis whereas Mr Crew wrote that the Claimant did not know any means of management or treatment and yet both gave him a 6. We have dealt with this at length above.

83. For Synthesis on scenario 1, Mr Kockelbergh gave a mark of 6, whereas Mr Crew marked it as a 5, the Claimant says with no explanation for the difference. As we have noted above, Mr Kockelbergh thought that the Claimant's reference to a ureteroscopy was incorrect. Mr Crew's written evidence, as again noted above, was that the Claimant did not understand radiation cystitis and its management, which he now accepts is incorrect, though he also referred to Mr Kockelbergh's notes about histopathology and healing problems.

84. For scenario 2, the Claimant says the discrepancy in the Synthesis domain (7 from Mr Kockelbergh, 6 from Mr Crew) is not explained. As again already noted, Mr Crew's evidence was that the absence of any notes indicates he regarded the candidate as safe.

85. The Claimant also says that if he had been given an examiner other than Mr Kockelbergh on this station, he would have received higher marks.

Station 2 (Mr Mills and Mr Shukla)

86. As we have noted, for scenario 1, Mr Mills gave the Claimant a mark of 5 for Knowledge and wrote (with underlining) "some prompting". The Claimant in his appeal (see page 759) says that the prompting never happened. Mr Mills told us that he cannot see why he would have underlined the word "prompting" if there had not been any. As also already noted, the marking descriptor at page 216 says in relation to "Quality of Response" that frequent prompting should be a 5, and minimal prompting a 6. Mr Mills told us that the Knowledge domain, tested at the start of each scenario, is the easy part, so that if a candidate stumbles at that point in relation to matters which should be at the forefront of their mind, that is "very disappointing". In his view, it is not just about the number of prompts required. As to why Mr Shukla gave a 6 for the same domain, he said he would expect him to make his own assessment.

87. For Application on scenario 1, the Claimant says there is no explanation for the difference between Mr Shukla's mark of 7 and Mr Mills' 6. As already noted, Mr Mills told us that the Application and Synthesis domains overlap. He accepts he ticked each comment he noted the Claimant as having made, but says it was all "basic urology". He adds that the Claimant's reference to a very well-known trial would not of itself have shown familiarity with literature meriting a mark of 7.

88. As for scenario 2, Mr Mills ticked what he recorded of the Claimant's comments in the Knowledge domain, giving him a 6 not a 7. He told us the Claimant "did well enough", understanding some features of the biopsy in question but being unsure about the use of a particular camera.

Station 3 (Ms Gall and Mr Woodward)

89. For Application, on scenario 1, Ms Gall made a reference on her mark sheet to "prompting" (it seems to have been assumed in oral evidence the phrase was "some prompting", but it in fact seems to be "after prompting"). The Claimant says that should have resulted in a 6 – he compares this to the mark at page 726 for example, from one of the unknown examiners, where it is said prompting was needed but a 6 is given. Ms Gall told us that notwithstanding the generic descriptor at page 216, even needing to be prompted once on something that

could lead to a significant problem for a patient, can lead to a 5: it is sound clinical judgment that matters in her view. Mr Shukla gave similar evidence.

90. The Claimant also says that Ms Gall (according to her witness statement) gave him 5s for two domains in scenario 1, for the same reason. Ms Gall accepted in oral evidence that each of the three categories should be marked separately, though she added that there is also a need to make a judgment overall.

91. The Claimant also points out that he had done scenario 1 in the 2018 exam – see page SB21 – for which he scored 666 ("solid performance" is what that examiner said) and 766. Ms Gall points out that follow up questions can differ even where the scenario is the same.

92. On scenario 2, the Claimant denies suggesting an invasive treatment as a first step (see his appeal document at page 772), though that is what Ms Gall's notes record, and she thinks it unlikely she and Mr Woodward made the same error.

Station 4 (Mr Summerton and Mr Mathur)

93. The Claimant says there is no explanation of the difference in scores of 666 and 777 for scenario 1. In the Application domain, the examiner giving a 6 referred to prompting being needed whereas Examiner 2 noted it as good. As we have noted, neither of the examiners from this station gave evidence at this Hearing.

Station 5 (Mr Hodgson and Mr Thomas)

94. The Claimant says there is also no explanation for the difference in mark for the Application domain on the second scenario, Mr Hodgson giving him a mark of 7, whereas Mr Thomas gave him a 6. He also says Mr Hodgson should have given him a 7 for Synthesis on scenario 2, because Mr Hodgson explicitly mentioned "MET" (medical expulsive therapy) on his mark sheet which is evidence that the Claimant mentioned research papers. The Claimant refers to page 742 where an examiner at another station scored him as a 7 for Synthesis, because he "quote[d] evidence". Mr Hodgson told us that MET is not a reference to literature, just to taking a tablet, but that the Claimant may well have referred to literature, in which case the 7 he gave him in the Application domain reflected that.

Station 6 (Mr Ravi and Mrs Webber)

95. The Claimant was marked 666 by both examiners for both scenarios. He says there were no negative comments and so the marks did not match the Respondent's marking descriptors.

Station 8 (Mr Napier-Hemy and Mr Morley)

96. For the Knowledge domain on scenario 1, one examiner said "safe and sensible", but gave the Claimant a 6; the Claimant says it should have been a 7. For the Synthesis domain, one examiner said "quotes evidence" and gave a 7, whilst the other examiner gave a mark of 6 noting, "limited progression with this scenario". Again, neither of the examiners from this station gave evidence.

97. The Claimant says he thus lost, unfairly, considerably more marks in total than the two marks by which he failed the Examination. The Respondent says he has been selective in his challenges to marks. It says:

97.1. As to prompting, it is not the number of prompts only that is relevant, but which questions require prompting.

97.2. Differences of one mark are not unusual. The number of marks and the fact of two examiners per station helps ensure fairness.

97.3. Where both examiners pass a candidate, no concern arises over differences in marks as far as the Respondent is concerned.

97.4. Quoting of evidence does not necessarily result in a good mark.

97.5. Notes are only an aide memoire. Any differences of two marks require justification.

97.6. Mr Ryan accepted that where there has been what one examiner at a station considers a basic omission, he would expect the other's mark to be aligned with that.

Direct discrimination

98. The complaint of direct discrimination is solely about the marks the Claimant was awarded. He has no issues with the scientific content of the exam or the marking scheme as such.

99. As the Claimant says, all examiners were aware of his name before the Examination. He says (AG38) that as a result they knew he was a non-UK national. He was inconsistent in his case as to which examiners were influenced by considerations of race in marking him, initially saying it was not every examiner, but then saying that all were and then – through Mr Grundy in closing submissions – saying that it was not all.

100. He cites the discrepancies in marking his exams referred to above as evidence as to why race was a feature in the outcome. He says the examiners wanted to show the superiority of UK training, because although the Respondent does not set the curriculum for the surgical training or indeed provide it, it is a "British institution". He says the Respondent was also punishing him for having raised fairness issues about his previous Examinations.

101. The Claimant also told us that an Egyptian colleague failed the Examination three times, essentially because of low marks given by Mr Kockelbergh, but then passed it when he got a UK passport. He does not blame Mr Kockelbergh for this, saying "it comes from the Respondent". Passports can be shown as proof of identity at an Examination but are not shown to examiners. We were not taken to any evidence regarding this colleague's marks nor was this point put to Mr Kockelbergh.

102. The Claimant also says the body language of examiners – frowning and looking away from him – showed him as they came to the station whether they

were going to “help him” or “let him down”, Mr Kockelbergh and Mr Woodward specifically falling in the latter category.

103. Finally, he points out that he passed on some scenarios in previous Examinations that he failed on his fourth attempt. He thinks it unlikely that this was based on his performance on the day.

104. The Respondent carries out an equality assessment after every set of examinations to report on exam reliability and how candidates in different groups perform. Its psychometric team also analyses how examiners fare across different protected characteristics. Its report is presented to the IQA committee. A more general report is also considered by the IQA and the Respondent’s main board and is sent to the GMC. None of the examiners in this case were shown to be scoring differently according to race.

105. Mr Ryan says at WR20 that the examiner body is so ethnically diverse that it makes meaningful biases “extremely unlikely”. The Respondent’s report on equality and diversity in examinations from 2009 to 2014 (see page 157ff) indicates that over that period 77.7% of examiners were White, with 13.4% Asian, whilst 85.9% of examiners had English as their first language. Mr Ryan adds (WR22) that an examiner’s behaviour is monitored by their colleague on the station, by an assessor and by the psychometric team’s subsequent analysis. He says at WR31 that “hawkishness or dovishness” is a potential behavioural issue: two or three examiners have been discussed in that context in the last 5 years. Half of the ten board chairs are from an ethnic minority (WR42). Mr Shukla accepted that it would be obvious the Claimant originated from outside of the UK on the basis of his name, colour and accent but was clear this would not indicate whether he was in training or not.

Indirect discrimination

106. The Claimant accepts that there has to be a figure chosen for passing the Examination. Mr Mills told us that if it were made two marks lower, for example, someone else would feel harshly treated. Mr Dickinson was unable to say why something like the Angoff technique was not applied to the Examination as well as to section 1, saying that was a “statistical question”.

107. The Claimant says that the Examination is failed disproportionately by people who do not have UK or EU nationality. According to the statistical analysis at page 591, for the Examinations from May 2016 to November 2019 there was a 92.8% mean pass rate for those in UK training posts and 47.6% for those not in UK training posts. The equivalent pass rates for the section 1 exam were 73.9% and 49.5% respectively. According to the analysis at page SB94, which we assume relates only to urology candidates, there were 599 surgical trainees who took the Examination in the period 2009-2019 and 358 candidates who were not in training. Of those numbers (the parties were agreed we could ignore the column headed, “Non-CCT trainees”):

107.1. The first attempt pass rate for trainees was 90% against 45% for those not in training.

107.2. The second attempt pass rate was 86% and 42% respectively.

107.3. The third attempt pass rate was 67% and 37% respectively.

107.4. The fourth attempt pass rate was 100% (2 out of 2) and 21% (3 out of 14) respectively. Thus, no candidate in training failed all four Examinations, whilst eleven not in training did.

108. Not all of those who failed the exam took it the next time, in other words there was what Mr Briggs described as some "attrition". We will come back to that point, and to page SB94 more generally in our Analysis.

109. At page 580ff is the Respondent's Equality and Diversity Report of 26 March 2021 assessing the period 2016 to 2020 across all specialties, not just urology. There were 2,792 candidates with English as their first language and 1,373 candidates with another first language (page 581). Pages 586 to 587 show, in cluster diagrams:

109.1. A strong correlation between BAME candidates not in training towards the CCT and failure of the Examination.

109.2. A less strong correlation between BAME candidates not in training towards the CCT and failure of section 1.

109.3. For White candidates taking the Examination, a strong correlation between those in training towards the CCT and those with higher scores.

109.4. For White candidates taking the section 1 exam, a less strong correlation between those in training towards the CCT training and passing it.

110. The part of this document at page 587 provides the same analysis for first language – English and "other" – again across all specialties, not just urology. Again, in both categories, there was a strong correlation between those in training towards CCT and passing the Examination, and a less strong correlation for passing section 1. The Respondent does not have data comparing pass rates for UK and non-UK nationals, though Mr Ryan agreed that it was fair and reasonable to suggest that first language analysis was the closest assessment to this available, as it is more likely that a person is not a UK national if their first language is not English. He also agreed there are some non-UK nationals whose first language is English, such as South Africans, but very few. He told us that around 10 to 15% of those who do UK surgical training in orthopaedics (his specialty) do not have English as their first language. The data at page 587 says that 86% of those candidates taking the section 1 examination and/or the Examination whose first language was English were in training (presumably meaning 14% of those whose first language was English were not in training). 28% of those taking the section 1 examination and/or the Examination whose first language was not English were in training, meaning that 72% of those taking the Examination whose first language was not English were not in training. As Mr Briggs noted in his written submissions, this means that of 2,792 candidates who spoke English as a first language (that total figure appears on page 581), 2,402 were in training and 390 were not in training, whilst of the 1,373 candidates (not 1,826 as suggested in Mr Briggs' submissions – again, see page 581) who did not speak English as a first language, 384 were in training and 989 were not in training (the figures are not exact but are nevertheless very close to exact). This can be represented as follows:

	In training	Not in training	Total
English language	2,402	390	2,792
Other language	384	989	1,373
Total	2,786	1,379	4,165

111. UK-trained candidates are given additional training, mentoring, and time off to prepare for the exams (AG48). The Claimant did not get that opportunity, as he was not in a training post. The Claimant says at AG49 that it is also clear that the Examination is intended to assess an applicant's ability to understand and perform the UK-training curriculum, rather than a candidate's ability in urology. This in his view plainly disadvantages those who have the relevant skillset but have not trained in the UK. He says that UK nationals are more likely to be trained in the UK than non-UK nationals. There was no data before us on this point, though some of the Respondent's witnesses accepted that there would be some correlation between not being a UK national and not having English as a first language.

112. The Respondent has known for some time that those in training posts pass the Examination at a much higher rate than those who are not. Mr Dickinson told us that concern about this has been expressed to the Royal Colleges and membership associations and that the Royal Colleges are seeking to provide opportunities for better training for those in non-training grades. Mr Ryan says at WR46 that most if not all doctors wanting to become NHS consultants apply for training posts, which are limited in number. There are more candidates for the CCT than places available. To get to the point of taking the Examination, nine assessments have to be passed. Trainees also attend a revision course for it. Those not in training posts do not get the same level of supervision, support or pre-testing. Whilst no statistics are available, the Claimant accepts that this would filter out certain individuals before the Examination was taken, though whilst also accepting that entrance into a surgical training post in the NHS is merit-based, in his view this does not necessarily reflect how participants progress thereafter. Mr Dickinson said that if there was any discrimination it was in how candidates were trained and not in the examination.

113. Mr Ryan described the requirement of the Examination to reach 576 marks as "all or nothing", saying that the Respondent had not thought about allowing unsuccessful candidates an opportunity to retake only the stations which were below the pass mark. He told us it is the same in most examinations and noted that a candidate can make up for failing one station by doing better on another. Mr Dickinson thinks it very unlikely a candidate would badly fail any individual station and still pass overall. Mr Ryan added that it is unlikely a candidate would go on to specialise as a surgeon in an area s/he failed in the Examination.

114. As already indicated, candidates are tested against the standard of a Day 1 consultant in the UK. The Claimant accepts that the GMC sets the standard required, but does not accept that this is a correct interpretation of it. Mr Mills told us that quality of care does not depend on geography, and that the reference to the UK is just about where the exam is done, but his fellow examiners did not agree. Ms Gall told us that the "day 1 consultant in the UK" standard is important because the system in other countries is very different in that some surgical staff abroad are "technicians", whereas in the UK there is a requirement to develop

treatment plans for example, involving the patient, their family and other experts. Mr Ryan's evidence was that the Examination is about providing consultants in the UK, hence this is the standard. Mr Dickinson similarly stated that this is an examination to qualify someone to work as a consultant in the NHS, a view shared by Mr Kockelbergh. He has only ever seen day 1 consultants in the UK and said that in other countries, the term "day 1 consultant" can signify something very different, for example in Germany they would not do surgery independently. Mr Ryan briefly referred in his evidence to an international examination which has the same content and examiners. Naturally that does not test against a standard of a day 1 urologist in the UK, but we heard nothing else about the international examination at all.

115. As for the limit of four attempts to pass the Examination, Mr Ryan told us that in the distant past the number of attempts was unlimited. It seems clear from pages SB75 and SB77 that in 2010 and 2014 respectively, the Respondent allowed fifth attempts and some limited re-entry. The latter meant going back to do section 1 again, though it is not clear that if this was passed a candidate could have another attempt at section 2. All those given these options were not trained in the UK. The Respondent's regulations for the Intercollegiate Specialty Examination in Urology 2015 (page 210ff) suggest that the limit of four attempts was introduced in 2015. The Claimant's case (see for example AG37) that the eleven candidates who failed at the fourth attempt (him and ten others) between 2009 and 2019 were non-UK nationals was not challenged.

116. The Claimant does not think there should be any limit to the number of attempts, as candidates can improve their performance over time. He suggested in cross-examination of the Respondent's witnesses that a candidate could be required to retake only the parts they failed, although Mr Mills said in response that this would make the Examination much easier to pass, the point being to assess the whole range of urological knowledge at a given point in time. The Claimant also suggested that a candidate could be allowed to pass overall but be required to do further training before practising in any area where the relevant station had been failed.

117. Mr Ryan told us that he thinks there was some research around ten to fifteen years ago which suggested that a candidate was unlikely to be successful after the third attempt, so that the demonstration of competence thereafter diminished and a candidate was likely to pass because of a combination of familiarity with the exam structure and questions and good fortune. Mr Ryan believes that the GMC informed the Royal Colleges of the research and that the Royal Colleges directed that the number of attempts be limited to four, on the basis of patient safety. He described three attempts as "the sweet spot", so that fixing it at four erred on the side of caution. It was decided by the President of the Royal Colleges, with the agreement of the GMC.

118. The Respondent says that the aim of the Examination, and thus the aim legitimising the PCPs, is ensuring, in the interests of patients and the general public, minimum levels of medical competence on the part of those who seek to become eligible for appointment as consultants. The Claimant is now working as a locum consultant on a fixed term contract with University Hospitals Coventry and Warwickshire NHS Trust, engaged in the full range of urology practice, including the more straightforward work in areas where his final Examination scores were lower (such as paediatric urology). He runs three clinics in his own

name each week, devising patient plans, and supervising diagnostic tests. He is in the operating theatre, without supervision, twice a week and is part of the on-call team. The only difference between his work and that of a substantive consultant is the length of the contract. He has a similar contract lined up in Leicester when the current one expires.

119. It is not necessary to pass Examination to work as a locum, though Mr Mills told us that a locum in the UK does not work independently, with everything s/he does being someone else's responsibility. Mr Ryan's evidence was somewhat different: he described locums as requiring "a bit of supervision" in his area of orthopaedics, being able to do most things themselves. Mr Kockelbergh, who is now a locum himself, effectively concurred, saying that the GMC makes a doctor responsible for their own actions, though if they are working in a clinic for someone else that person may have ultimate responsibility.

Time limits

120. When asked about it in evidence, the Claimant said almost nothing about why the complaint based on the fourth PCP had not been presented before the Tribunal permitted the amendment during this Hearing. He said he did not know why it had not been presented before; he thought it had been included in his "submission" at some point. He did not understand PCPs and the difference between direct and indirect discrimination. He got informal legal advice from a friend of a friend when he submitted his appeal to the Respondent against the outcome of the fourth Examination in late 2019 and first instructed solicitors in April or May 2020.

Witnesses

121. We have noted already that the Respondent did not call as witnesses Messrs Thomas or Woodward, though both gave statements. We have also noted those stations in respect of which no witness was called and no statement taken. The Respondent provided no specific explanation for the absence of these witnesses other than that they are engaged in busy clinical practices at time of junior doctors' strike (mentioned by Mr Briggs earlier in the Hearing and in closing submissions) and that it was a "litigation decision", including taking the view that it did not add anything to the case and would have been disproportionate to call sixteen examiners to cover substantially the same ground.

Law

Burden of proof

122. Section 136 of the Act provides as follows:

(1) This section applies to any proceedings relating to a contravention of this Act.

(2) If there are facts from which the court [which includes employment tribunals] *could decide, in the absence of any other explanation, that a person (A) contravened the provision concerned, the court must hold that the contravention occurred.*

(3) But subsection (2) does not apply if A shows that A did not contravene the provision".

123. Direct evidence of discrimination is rare and tribunals frequently have to consider whether it is possible to infer unlawful conduct from all the material facts. This has led to the adoption of a two-stage test, the workings of which were described in the annex to the Court of Appeal's judgment in **Wong v Igen Ltd [2005] ICR 931,** updating and modifying the guidance that had been given by the EAT in **Barton v Investec Henderson Crosthwaite Securities Ltd [2003] ICR 1205.** The Claimant bears the initial burden of proof. The Court of Appeal held in **Ayodele v Citylink Limited and anor [2017] EWCA Civ. 1913** that "there is nothing unfair about requiring that a Claimant should bear the burden of proof at the first stage. If s/he can discharge that burden (which is one only of showing that there is a prima facie case that the reason for the Respondent's act was a discriminatory one) then the claim will succeed unless the Respondent can discharge the burden placed on it at the second stage".

124. At the first stage, the tribunal does not have to reach a definitive determination that there are facts which would lead it to the conclusion that there was an unlawful act. Instead, it is looking at the primary facts to see what inferences of secondary fact could be drawn from them**.** As was held in **Madarassy v Nomura International plc [2007] IRLR 246**, "could conclude" refers to what a reasonable tribunal could properly conclude from all of the evidence before it, including evidence as to whether the acts complained of occurred at all. In considering what inferences or conclusions can thus be drawn, the tribunal must assume that there is no adequate explanation for those facts.

125. If the burden of proof shifts to the Respondent, it is then for it to prove that it did not commit, or as the case may be, is not to be treated as having committed, the allegedly discriminatory act. To discharge that burden, it is necessary for the Respondent to prove that the treatment was in no sense whatsoever on the grounds of race. That would require that the explanation is adequate to discharge the burden of proof on the balance of probabilities, for which a tribunal would normally expect cogent evidence.

126. All of the above having been said, the courts have warned tribunals against getting bogged down in issues related to the burden of proof – **Hewage v Grampian Health Board [2012] ICR 1054**. In some cases of direct discrimination, it may be appropriate for the tribunal simply to focus on the reason given by the Respondent for its acts or omissions and if it is satisfied that this discloses no discrimination, then it need not go through the exercise of considering whether the other evidence, in the absence of a satisfactory explanation, would have been capable of amounting to a prima facie case of discrimination.

127. The implications of **Hewage** were considered by the EAT in **Field v Steve Pye and Co (KL) Ltd and others [2022] EAT 68**. The EAT said that where there is significant evidence that could establish that there has been discrimination, it cannot be ignored. In such a case, where a tribunal moves straight to the "reason why" question it could only do so on the basis that it has assumed the claimant has passed the stage one threshold, so that the burden was now upon the respondent in the way described above. The EAT went on to say that if at the end of the hearing the tribunal concludes that there is nothing that can suggest that discrimination has occurred and the respondent has established a non-

discriminatory reason for the impugned treatment, there would be no error of law in just answering the "reason why" question, but in fact the complaint would fail at the first stage. If having heard all of the evidence the tribunal concludes that there is some evidence that could indicate discrimination, but nonetheless is fully convinced that the impugned treatment was in no sense whatsoever because of the protected characteristic, it is permissible to reach a conclusion at the second stage only, but there is much to be said for properly grappling with the evidence and deciding whether it is sufficient to shift the burden of proof. Particular care should be taken if the reason for moving to the second stage is to avoid the effort of analysing evidence that could be relevant to whether the burden of proof should have shifted at the first stage.

Section 53 of the Act

128. It is section 53 of the Act which renders discrimination by qualifications bodies unlawful. It provides, so far as relevant to the issues we had to decide:

(1) A qualifications body (A) must not discriminate against a person (B) –

(a) in the arrangements A makes for deciding upon whom to confer a relevant qualification …

(c) by not conferring a relevant qualification on B.

Direct discrimination

129. Section 13 of the Act provides, again so far as relevant:

(1) A person (A) discriminates against another (B) if, because of a protected characteristic, A treats B less favourably than A treats or would treat others.

130. The protected characteristic relied upon in this case is race. Section 23 provides, as far as relevant:

(1) On a comparison of cases for the purposes of section 13 … there must be no material difference between the circumstances relating to each case".

131. The Tribunal must therefore consider whether one of the sub-paragraphs of section 53(1) is satisfied, whether there has been less favourable treatment than would have been afforded to a hypothetical comparator, and whether this was because of the Claimant's race. A difference in nationality and in treatment is not sufficient – **Madarassy**; there must be something more, though Mr Grundy correctly submitted this "need not be a great deal", referring us to **Deman v The Commission for Equality and Human Rights and others [2010] EWCA Civ. 1279**.

132. The fundamental question in a direct discrimination complaint is the reason why the Claimant was treated as he was. As Lord Nicholls said in **Nagarajan v London Regional Transport [1999] IRLR 572** "this is the crucial question". Race being part of the circumstances or context leading up to the alleged act of discrimination is insufficient. Whilst the question of less favourable treatment plainly cannot be ignored, it is acceptable for tribunals to address the reason why question first, as in many cases that will answer the question of

whether there has been less favourable treatment – see **Shamoon v Chief Constable of the Royal Ulster Constabulary [2003] ICR 337, HL** and **Stockton on Tees Borough Council v Aylott [2010] ICR 1278, CA**. This is how the Claimant pursued his case, as confirmed by Mr Grundy in closing.

133. In most cases – such as **Nagarajan** – the act complained of is not in itself discriminatory but is rendered discriminatory by the mental processes (conscious or otherwise) which led the alleged discriminator to act as they did. Establishing the decision-maker's mental processes is not always easy. What tribunals must do is draw appropriate inferences from the conduct of the alleged discriminator and the surrounding circumstances. In determining why the alleged discriminator acted as they did, the Tribunal does not have to be satisfied that the protected characteristic was the only or main reason for the treatment. It is enough for the protected characteristic to be significant in the sense of being more than trivial (again, **Nagarajan** and **Igen**). Unreasonable behaviour of itself is not evidence of discrimination – **Bahl v The Law Society [2004] IRLR 799** – though the Court of Appeal said in **Anya v University of Oxford and anor [2001] ICR 847** that it may be evidence supporting an inference of discrimination if there is nothing else to explain it.

Indirect discrimination

134. Section 19 of the Act provides that indirect discrimination occurs when a person (A) applies to another (B) a PCP that is discriminatory in relation to a relevant protected characteristic of B's. This is the case when, according to section 19(2):

(a) A applies, or would apply, [the PCP] to persons with whom B does not share the [relevant protected] characteristic [here, race],

(b) it puts, or would put, persons with whom B shares the characteristic at a particular disadvantage when compared with persons with whom B does not share it,

(c) it puts, or would put, B at that disadvantage, and

(d) A cannot show it to be a proportionate means of achieving a legitimate aim.

Section 23 must be taken into account when determining the question of "group disadvantage" posed by section 19(2)(b) – see above.

135. The Claimant bears the burden of proof in respect of the first three steps in section 19(2). In **Ishola v Transport for London [2020] ICR 1204, CA**, the Court of Appeal held that the words, "provision, criterion or practice" are broad and overlapping and not to be narrowly construed or unjustifiably limited in their application. The function of a PCP is to identify what it is about the employer's management of the employee [in that case] or its operation which causes the particular disadvantage. The Court stated, "to test whether the PCP is discriminatory or not, it must be capable of being applied to others because the comparison of disadvantage caused by it has to be made by reference to a comparator to whom the alleged PCP would also apply". The Court added, "however widely and purposively the concept of a PCP is to be interpreted, it does not apply to every act of unfair treatment of a particular employee. That is

not the mischief which the concept of indirect discrimination … [is] intended to address … in context, all three words carry the connotation of a state of affairs … indicating how similar cases are generally treated".

136. In relation to the second and third steps, we considered the decisions in **Essop v Home Office; Naeem v Secretary of State for Justice [2017] UKSC 27** and **Allen v Primark Stores Limited [2022] EAT 57**, in addition to **Ryan v South-West Ambulance Service NHS Trust [2021] ICR 555** and drew from them the following principles:

136.1. It is for the Tribunal to determine the pool which will be used to determine whether section 19(2)(b) ("group disadvantage") is satisfied. See also **Grundy v British Airways plc [2008] IRLR 74** in which the Court of Appeal held that provided it tested the allegation in a suitable pool, a tribunal could not be said to have erred in law even if a different pool with a different outcome could legitimately have been chosen.

136.2. The EHRC Code states, "In general, the pool should consist of the group which the provision, criterion or practice affects (or would affect) either positively or negatively, while excluding workers who are not affected by it, either positively or negatively" – paragraph 4.18.

136.3. The pool must test the particular discrimination complained of.

136.4. More than one pool may be perfectly logical and therefore permissible.

136.5. Section 23 requires that those in the pool must not be in materially different circumstances.

136.6. The required link to establish indirect discrimination is not between the protected characteristic and the group and individual disadvantage but between the PCP and the disadvantage.

136.7. It is not necessary to establish why the PCP causes the disadvantage – the reasons for such disadvantages, if identifiable, can be many and varied and are often innocuous.

136.8. Not every member of the disadvantaged group needs to suffer the disadvantage.

136.9. The individual and group disadvantage must correspond.

137. In **Pendleton v Derbyshire County Council [2016] IRLR 580** the EAT did not read "particular disadvantage" for these purposes as requiring any particular level or threshold of disadvantage. The term was "apt to cover any disadvantage". In brief, the question is whether more people of non-UK nationality in the selected pool experience the disadvantage than those of UK nationality. There will need to be some basis on which to conclude that this is the case, though in **Homer v Chief Constable of West Yorkshire [2012] ICR 704** Baroness Hale noted that "the new formulation [in the Act] was not intended to make it more difficult to establish indirect discrimination: quite the reverse … It was intended to do away with the need for statistical comparisons where no statistics might exist. It was intended to do away with the complexities involved

in identifying those who could comply [with the PCP] and those who could not and how great the disparity had to be. Now all that is needed is a particular disadvantage when compared with other people who do not share the characteristic in question. It was not intended to lead us to ignore the fact that certain protected characteristics are more likely to be associated with particular disadvantages".

Justification

138. We draw the following principles from the relevant case law concerned with whether the PCP is a proportionate means of achieving a legitimate aim (justification for short):

138.1. The burden of establishing this defence is on the Respondent.

138.2. The Tribunal must undertake a fair and detailed assessment of the Respondent's business needs and working practices, making clear findings on why the aims relied upon were legitimate, and whether the steps taken to achieve those aims were appropriate and necessary.

138.3. What the Respondent does must be an appropriate means of achieving the legitimate aims and a reasonably necessary means of doing so. In **Homer** it was said, approving Mummery LJ **in R (Elias) v Secretary of State for Defence [2006] EWCA Civ 1293**, and mirroring the decision in **Bilka-Kaufhaus GmbH v Weber Von Hartz [1987] ICR 110** referred to in Mr Grundy's written submissions, that what is required is: first, a real need on the part of the Respondent; secondly, that what it did was appropriate – that is rationally connected – to achieving its objectives; and thirdly, that it was no more than was necessary to that end.

138.4. In **Hardy & Hansons plc v Lax [2005] ICR 1565** it was said that part of the assessment of justification entails a comparison of the impact upon the affected person as against the importance of the aim to the employer. It is not enough that a reasonable employer might think the treatment justified.
The Tribunal itself has to weigh the real needs of the Respondent, against the discriminatory effects of the aim. A measure may be appropriate to achieving the aim but go further than is (reasonably) necessary in order to do so and thus be disproportionate.

138.5. It is also appropriate to ask whether a lesser measure could have achieved the employer's aim – **Naeem**.

138.6. In summary, the Respondent's aims must reflect a real business need; the Respondent's actions must contribute to achieving it; and this must be assessed objectively, regardless of what the Respondent considered at the time. Proportionality is about considering not whether the Respondent had no alternative course of action, but whether what it did was reasonably necessary to achieving the aim.

Time limits

139. Section 123(1) of the Act provides that proceedings on a complaint under Section 120 may not be brought after the end of the period of 3 months starting

with the date of the act to which the complaint relates, or such other period as the Employment Tribunal thinks just and equitable. It was agreed that this was the only time limit issue in this case.

140. The provision for extending time where it is just and equitable to do so gives to tribunals wider scope than the test of reasonable practicability which applies for example in unfair dismissal cases. Nevertheless, there is no presumption that time will be extended – **Robertson v Bexley Community Centre (trading as Leisure Link) [2003] IRLR 434**, though extending time does not require exceptional circumstances. In **British Coal Corporation v Keeble [1997] IRLR 336**, it was held that similar considerations arise in this context as would be relevant under the Limitation Act 1980, namely the prejudice which each party would suffer as a result of the tribunal granting or refusing an extension, and all the other circumstances, in particular: (a) the length of and reasons for the delay; (b) the extent to which the cogency of the evidence is likely to be affected by the delay; (c) the extent to which the party sued had co-operated with any requests for information; (d) the promptness with which the claimant acted once he or she knew of the facts giving rise to the cause of action; and (e) the steps taken by the claimant to obtain appropriate professional advice once he or she knew of the possibility of taking action.

141. In **Abertawe Bro Morgannwg University Local Health Board v Morgan [2018] EWCA Civ. 640** Leggatt LJ in the Court of Appeal said that Parliament has given tribunals "the widest possible discretion" in deciding whether to extend time in discrimination cases. Notwithstanding **Keeble** there is no list of factors which a tribunal must have regard to, though the length of and reasons for delay, and whether delay prejudices a respondent for example by preventing or inhibiting it from investigating the claim whilst matters were fresh, will almost always be relevant factors. He said that there is no reason to read into the statutory language any requirement that the Tribunal must be satisfied that there are good reasons for the delay, let alone that time cannot be extended in the absence of an explanation of delay from the Claimant. At most, he said, whether any explanation or reason is offered and the nature of them are relevant matters to which the Tribunal should have regard.

142. **Adedeji v University Hospitals Birmingham NHS Foundation Trust [2021] EWCA Civ 23** approved **Morgan**. In that case, a race discrimination complaint was dismissed when three days out of time. Two key factors in the Court of Appeal upholding the decision were that the events in question occurred a long time beforehand, and that the claimant in that case was a very educated man who had taken legal advice.

Other

143. Mr Grundy referred us to the EAT's recent decision in **General Medical Council v Dr O M A Karim [2023] EAT 87**. That seems to us very much a decision on its own facts, for example where the employment tribunal did not make clear whether a particular complaint had succeeded or not. It made the additional point that where statistics are utilised as part of a claimant's case as to why the burden of proof should pass to the respondent, it is incumbent on the tribunal to grapple with the respondent's evidence, if any, in relation to those statistics.

Analysis

Indirect discrimination

144. I turn now to our conclusions, and for reasons which will hopefully become clear, deal first with the complaints of indirect discrimination.

Did the Respondent have one or more PCPs?

145. The Respondent did not dispute that it applied the PCPs set out in the list of issues above at paragraphs 6.1, 6.3 and 6.4 (referred to hereafter as PCP1, PCP3 and PCP4 respectively). It did not accept however that it applied the PCP set out in the list of issues at paragraph 6.2 above (referred to hereafter as PCP2). That was therefore the first issue we were required to determine.

146. The alleged PCP was formulated as follows: "marking down candidates who were not in UK/EU training posts and/or marking down non-UK national candidates and/or minority ethnic candidates". As I indicated to Mr Grundy during his closing submissions, we could not see how the part of this formulation falling after the words "and/or" could possibly be a PCP, as what those words describe was by its nature incapable of applying to candidates for the Examination who were UK nationals and/or not minority ethnic candidates. That part of the PCP is in fact a complaint of direct discrimination, in that it alleges that the Respondent marked down (which it is agreed means marked wrongly – see below) non-UK nationals and ethnic minority candidates and did not mark down UK nationals and candidates who were not from an ethnic minority – in other words, less favourable treatment because of race. There could therefore as a matter of law, and of logic, be no such PCP.

147. Separately, and as Mr Grundy accepted in his closing submissions, there was nothing whatsoever in the presentation of the Claimant's case which touched on the position of candidates who were or were not in EU training posts. This left us to consider whether the Respondent applied a PCP of "marking down candidates who were not in UK training posts", which we accept in principle could be a PCP given that there were candidates of both UK nationality and non-UK nationality who were not in such posts.

148. As just indicated, Mr Grundy accepted that the alleged PCP referred to a practice of marking candidates wrongly in the November 2019 Examination: it could plainly have no other meaning than giving candidates lower marks than their answers merited. The question for us to determine was whether the Claimant had established that the Respondent did in fact apply such a PCP.

149. As already noted, there was no evidence presented to us of any candidate's marks other than the Claimant's, specifically in this context a candidate not in a UK training post, whether they were a UK national or a non-UK national. We will come separately to the question of whether the Claimant himself was marked down when we deal with his direct discrimination complaint, but even if he was, noting the decision in **Ishola** and given the particular nature of the PCP in question, in our judgment his marks alone were not sufficient to show that there was, or would be, to any extent a wider practice of marking down such candidates.

150. The Claimant relied in addition on statistical information, in particular information showing that those not in UK training posts did not do so well in the Examination as those who were in such posts. We will come to those statistics below, but the point to note here is that the fact that non-trainees fared worse overall is not evidence that they were wrongly or unfairly marked. We are conscious of course, as the Supreme Court in **Essop** made clear, that it is not necessary for a claimant to be able to show the reason why persons sharing their protected characteristic were put to a particular disadvantage, but in this case that is precisely what the Claimant would need to establish to get the complaint based on this PCP off the ground. In other words, it is not enough for the purposes of showing that the Respondent had this particular PCP to say that non-trainees got generally lower marks than trainees; it would have to be shown that this was because they were marked wrongly. We were not taken to any such evidence, beyond the Claimant's assessment of his own performance and marking. The possible exception was the Claimant's reference to his colleague who passed the Examination once he got a British passport, but first, even if examiners were aware of his change of passport status that would not indicate to them whether he was a trainee or not, and secondly his case was not put to the Respondent's witnesses. Further and in any event, we had no evidence that he had previously been wrongly marked. For those reasons we attach no weight to that evidence.

151. We would further add that whilst some of the examiners knew that the Claimant, and some other candidates, were non-UK nationals (see our conclusions below when dealing with the complaint of direct discrimination), we are satisfied that none of them had knowledge of which candidates were in UK training posts and which were not.

152. We note too, as Mr Briggs pointed out in his submissions, the following:

152.1. Over the period from 2016 to 2020, 384 out of 2,786 candidates in UK training posts, that is 14% (not a third as we understood Mr Briggs to suggest), did not have English as their first language and 390 out of 1,379, that is 28%, of candidates who were not in UK training posts spoke English as their first language. This means that examiners who consciously or unconsciously sought to mark down those not in UK training posts based on an assumption founded on first language would have a significant chance of getting it wrong. The data on page 587 suggests that those who speak English as a first language who are trainees do better than those who speak English as a first language who are not.

152.2. If examiners were marking down those not in UK training posts based on assumptions made from perceived ethnicity, one would expect to see similar pass rates amongst trainees and non-trainees within each ethnic group, whereas as we have already indicated, the data at page 586 indicates both that BAME candidates in training did better than BAME candidates not in training and that White candidates in training did better than White candidates not in training.

153. Finally, as already indicated and unusually for a complaint of indirect discrimination, this particular PCP requires a conscious or unconscious decision on the part of examiners to mark down non-trainees. We cannot see that the examiners had any reason to do so, given especially that the Respondent has for some time been concerned about the lower pass rates amongst non-trainees and has reported this concern to the Royal Colleges.

154. For all of these reasons, the Claimant has not established that the Respondent applied PCP2. His complaint of indirect discrimination depending on that PCP fails on that basis.

Disadvantage

155. The next question to be determined was whether any of the other PCPs put persons of non-UK nationality at a particular disadvantage compared to persons of UK nationality, in that they were less likely to pass the Examination. We deal first with PCPs 1 and 3, which we were satisfied could be taken together, concerned as they both were with the question of the standard that had to be achieved to pass. As already noted, it was agreed when setting out the list of issues at the start of the Hearing that the Respondent applied the PCPs to all candidates. That said, neither party rigidly adopted that starting point when identifying the pool of candidates within which particular disadvantage should be considered, the Respondent as to its principal position and the Claimant as to his secondary, alternative position. Neither party suggested that the other was prevented from framing their case differently to that provided by the wording of the list of issues as set out above.

156. Other than in relation to the Claimant's alternative basis for establishing particular disadvantage (see below), there was no direct evidence before us as to the respective achievement rates of UK nationals and non-UK nationals and therefore in respect of PCPs 1 and 3, the Claimant's principal case relied on the data at page 591. This, as we have already explained, shows significant differential attainment between those in UK training posts and those who were not, across all specialties. Mr Grundy's argument proceeded in four steps. First, as he highlighted in his written submissions, Mr Ryan accepted that those who did not have English as their first language were more likely to be non-UK nationals and vice versa. Secondly, as is evident from the data on page 587, of candidates between 2016 and 2020 who spoke English as their first language, 86% were in training, whereas only 28% of those who did not speak English as a first language were in training. Thirdly therefore, and broadly speaking, if a candidate was in training, they were more likely to be a UK national. One could therefore conclude on the basis of this broad analysis, fourthly, that non-UK nationals (more likely not to be in training) were much more likely to fail the Examination.

157. The Respondent's case was that this analysis, based on putting all candidates for 2016 to 2020 (or any period) into the pool for comparison, creates a highly misleading picture, on the simple basis that it is universally accepted that having a training post made a candidate more likely to pass the Examination. We have set out a number of reasons for that submission in our findings of fact, but in summary the evidence shows the following:

157.1. A training post is essentially supernumerary, with a focus on training towards obtaining the CCT, whereas those in non-training posts will have substantial non-training contractual responsibilities taking up the lion's share of their time.

157.2. Those in training posts thus have a far greater opportunity, over several years, to attend courses and conferences, including courses specifically designed to prepare them for the Examination – we recall Mr Woodward's

reference to the importance of “exam technique”, which was plainly another way of saying that trainees are better prepared to pass.

157.3. They have better access to supervision and mentoring from more senior clinicians and to peer support.

157.4. They are subject to regular work-based assessments.

157.5. One could also add that securing a training post in the first place is recognised to be a competitive process, such that to some extent at least, there appears to be a degree of academic strength built into the trainee cohort.

157.6. The Claimant himself accepted that training made a big difference, and was keen to highlight that he did not get the opportunities afforded to trainee candidates in respect of preparation and support for the Examination.

157.7. As the data for 2016 to 2020 at page 586 indicates, BAME candidates not in training and White candidates not in training were less likely to pass. The data at page 587 makes the same point in relation to language. The section 1 results show the same picture, albeit in a less pronounced way.

158. Whilst again keeping in mind the warning of the Supreme Court in **Essop/Naeem** that it is not necessary to establish the reason for a particular disadvantage, we accept Mr Briggs’ submission that it is crucial to recognise that the Respondent is not, and has not ever been, responsible for who becomes a trainee or for the training they are given, nor indeed has it ever been responsible for the absence of equivalent training and support for non-trainees. It is responsible for the Examination. As he submitted therefore, any differential in achievement between trainees and non-trainees is locked in before the Examination is taken. As the Supreme Court also said in **Essop/Naeem**, the pool must suitably test the discrimination complained of. We accept Mr Briggs’ submission that if we assessed the particular disadvantage (of being less likely to pass the Examination) within a pool of all candidates, we would be at considerable risk of testing the question of training discrimination rather than race discrimination. This is not to fall into a stereotypical assumption, asserted by the Claimant to exist amongst some or all examiners, that UK training is superior to what can be secured elsewhere. The Examination was tailored to the UK trainee curriculum and was part of the process of fitting candidates to be registered with the GMC as suitable for consultant posts in the UK. It is simply the context in which the Examination took place and is how the witnesses before us – including the Claimant (see paragraph 157.6 above) – viewed it.

159. In view of the above, Section 23 prevents us from assessing the claimed discrimination by adopting the Claimant’s principal approach of comparing trainees and non-trainees, because in our judgment this would involve comparing candidates between whom there was a highly material difference. Mr Grundy’s reply to that point during closing submissions was to say that the Claimant’s greater practical experience compared to those in training posts meant that there was no material difference. Whilst of course the Claimant’s experience in urological practice is substantial, the difficulty with that argument is that it is entirely unknown to us whether he is typical of the non-trainee group and even if he was typical, that would not address the obvious, accepted, point that trainees

are better prepared for passing the Examination, something for which the Respondent is not responsible.

160. The Claimant simply has not provided the evidence to show any comparison between the alleged advantaged (UK national) and disadvantaged (non-UK national) groups within what we regard as the appropriate pool, that is of those candidates not in training posts. He has not therefore satisfied the burden that is on him to establish a particular disadvantage for non-UK nationals in this way.

161. Mr Grundy said in closing that as an alternative argument on particular disadvantage, the Claimant relies on the fact that the eleven candidates who failed the Examination four times between 2009 and 2019 were all non-UK nationals. As we have said, the Claimant's assertion about their nationality was not challenged by the Respondent, and Mr Briggs accepted in closing submissions that it was a logical inference. We therefore turn next to assess that argument.

162. The document where this information is set out is that at page SB94 which, as we have noted, seems to be a ten-year assessment of urology candidates only. Mr Briggs submitted that the eleven candidates is too small a number to be properly representative of the true position. He made that submission in the context of the effect of PCP4, which we will come to, but in our judgment what he says is clearly correct when seeking to assess the effects of PCPs 1 and 3, given that, at this point, it is the effects of the pass mark and of the assessment against the standard of a day 1 consultant in the UK, and not the question of the maximum number of attempts to pass, which is in view. A fairer, more representative picture of the effects of the standard required to pass (whether PCP1 or PCP3) will be observed by looking at, for consistency with our analysis above, the out of training candidates as a whole.

163. We make the following observations about page SB94:

163.1. We noted during our deliberations that there appear to be some problems with the document in that, for example, there are said to have been 358 out of training candidates, but only 206 took the Examination first time. It is wholly unclear what happened to the remaining 152. We therefore treat the document with a degree of caution.

163.2. Nevertheless, as we have said, it is accepted that no UK nationals in this group failed four times and that eleven non-UK nationals did.

163.3. It is however unknown know how many of the 358 out of training candidates were UK nationals and how many were non-UK nationals, so that no overall comparison between UK national and non-UK national pass rates can be ascertained, nor can that exercise be done for failures at the first, second and third attempts.

164. The eleven non-UK nationals who failed at the fourth attempt represent a very small actual number of candidates. Moreover, as indicated above, it is not known how many UK and non-UK nationals respectively passed on the first, second or third attempts, so that how the pass rate and Examination standard was operating at <u>those</u> points is wholly unclear. In our judgment, it would be important to know how these PCPs were operating on those occasions to

properly test their effect. For these reasons, we are not satisfied that the Claimant has established on his alternative case a particular disadvantage for non-UK nationals occasioned by the overall pass mark (PCP1) or the requirement to meet the standard of a day 1 consultant in the UK (PCP3). The complaints based on those PCPs fail on that basis.

165. Turning to PCP4, the Respondent's starting position seemed to be that as a matter of principle there can have been no particular disadvantage for non-UK nationals in not being permitted more than four attempts to pass the Examination because it is not known how many of them would have taken the Examination on further occasions nor how many would have passed it. We do not accept that this means that only being able to take the Examination four times could not lead to any disadvantage in principle. We think that in principle it could. As Mr Grundy submitted, the disadvantage could be said to be not having the opportunity and thus failing the Examination overall.

166. We are not satisfied however that the data at page SB94 establishes a particular disadvantage for non-UK nationals as a result of this PCP. If one assumes that 28% of the 358 out of training candidates were UK nationals (adopting the percentage from the table at paragraph 110 above), that means there were 100 in total, and thus 258 non-UK national candidates. The eleven who failed four times thus represent 4.3% of the non-UK nationals within this group overall. This compares to 0% of UK nationals in the group overall who failed four times. Focusing on advantage rather than disadvantage, this means that 95.7% of non-UK nationals passed within the allotted attempts compared to 100% of UK nationals who did so. These are small percentage differences and do not represent a large actual number of candidates.

167. Furthermore, it is also relevant to consider the drop-out rates at earlier stages – what Mr Briggs described as "attrition". Noting our small caution about page SB94 generally:

167.1. At the first attempt, 112 out of training candidates failed. Only 98 took the second attempt, meaning that 14 dropped out altogether.

167.2. At the second attempt, 56 candidates failed. Only 51 took the third attempt, meaning that 5 dropped out altogether.

167.3. At the third attempt, 32 candidates failed. Only 14 took the fourth attempt, meaning that 18 dropped out altogether.

167.4. At the fourth attempt, 11 out of 14 failed, including the Claimant.

168. The point of this analysis is to say that there is no way of knowing how many candidates who failed at the fourth attempt would have wished to continue to a further attempt. That is in our view a relevant consideration in assessing whether particular disadvantage flowing from the PCP has been made out, because there was no such disadvantage at all for those who would have decided not to take the Examination on a further occasion even if they had been permitted to. The Respondent did not have a practice of stopping a candidate taking a fourth Examination after failing the third and yet eighteen candidates decided not to do so, an attrition rate at that stage of over 50%. Accordingly, the number of non-UK national candidates who in practice were disadvantaged by the PCP, as

opposed to those who would have dropped out of their own accord or for other reasons had there been another opportunity, is likely to have been substantially fewer than eleven, very likely no more than a similar 50% thereof (and probably less than that), so that the actual extent of the group disadvantage would be in the order of 2% compared to the 0% for UK nationals. In other words, 98% of non-UK nationals would have passed within the allotted attempts or decided not to proceed further, compared to 100% of UK nationals. We do not regard that as establishing a particular disadvantage for non-UK nationals, again particularly when having regard to the actual small number of candidates adversely affected by the PCP.

169. The Claimant has thus not met the burden on him to show that non-UK nationals were put to a particular disadvantage compared to UK nationals by any of the PCPs, meaning that nothing need be said about particular disadvantage for the Claimant himself. We note in addition his comment that he would have passed the Examination if he had been marked by different examiners. That seems to us to rather undermine his case that the pass mark, the adoption of a day 1 consultant in the UK standard, and even the maximum of four attempts, put him at a particular disadvantage compared to UK nationals. It confirms that in fact his case in his own mind was one of direct discrimination, to which we will come shortly. Although the complaints of indirect discrimination fail on the basis we have outlined, we did go on to briefly consider the questions of justification and time limits and would have held that the complaints failed on those grounds also. We deal briefly with each in turn.

Justification and time limits

170. We adopt the term “justification” as short-hand for the statutory requirement on the Respondent, had it been necessary, to show that the PCPs were a proportionate means of achieving a legitimate aim. The Claimant did not contest that needing to keep patients and the public safe and give them confidence in consultants was a legitimate aim. The question would thus have been whether the PCPs were a proportionate means of achieving it.

171. In relation to PCP1, the overall pass mark, the Respondent followed a rigorous calibration process at the start of each diet (including in November 2019), involving multiple experienced consultants, to determine the sort of answer required to secure a pass mark in relation to each question, so that as Mr Dickinson put it, the pass mark is fixed at a level which the consultants would regard as representative of safe practice. It seems to us abundantly clear therefore that the pass mark reflected a real need on the part of the Respondent, required of it by the Royal Colleges (and, it might also be said, the GMC) to make its contribution through the Examination process to securing the critical aim of patient and public safety. As Mr Briggs submitted, the Respondent does not have to show that in and of itself its PCP wholly secured achievement of the legitimate aim. It has to be seen as part of a series of responsibilities shared by the ISCP, the JCST, the GMC and indeed NHS trusts as employers.

172. It seems equally clear that fixing the pass mark through the calibration process provided a clear rational connection to achieving that aim. As to whether it was no more than was necessary to achieve the aim and whether a lesser measure might have done so, in our judgment the Respondent would have

established that the PCP was a proportionate means of achieving the aim against those principles established by the case law referred to above:

172.1. The Examination was recognisably not easy to pass. That is of course how it should be, given the aim in question, but we think it can also be fairly said that the pass mark was not set too high. It is somewhat deceptive to characterise it as "75%", when the lowest mark that could be awarded was 50%.

172.2. Not achieving the pass mark did not act as a wholesale bar to securing entry on to the GMC's Specialty Register, given the opportunity for a total of four attempts and the availability of the CESR route. We accept that the latter was very difficult to achieve, though we also note that it would have been open to the Claimant – and open to any experienced practitioner in his position – to apply for a training post, which can lead to a much-shortened pathway to obtaining the CCT.

172.3. As Mr Briggs submitted, not achieving the pass mark was not a wholesale bar to a fulfilling and successful medical career, as the Claimant himself demonstrates.

172.4. As the Claimant points out, a candidate did not have to pass at all eight stations in order to reach 576 marks overall. His point was that this makes the pass mark unjustifiable, but in fact it seems to us to have contributed to the proportionality of the mark, in that there were multiple ways of achieving it, at the hands of multiple different examiners.

172.5. As to the point that this means a candidate could practice in an area where they have failed the Examination, because they have passed overall, it is true that the Respondent could have constructed the Examination differently, for example fixing the pass mark at six out of eight stations or requiring that candidates pass eight out of eight stations which would be the more natural concomitant of the Claimant's argument. The former might well have benefited the Claimant (the latter would not), but this does not mean that the requirement to pass by achieving an overall mark, designed to test and reflect overall urological knowledge and judgment at a particular point in time, was a disproportionate means of securing the desired end, not least given that the Respondent is not aware that what the Claimant posits has actually happened in practice and Mr Ryan's evidence that a candidate is unlikely to specialise in surgical practice in an area where they failed the Examination.

172.6. That in our view is also part of the answer to the Claimant's case that his work as a locum means that the Examination requirements went too far. Mr Grundy said in submissions that the point of this argument was that how the pass mark was set should have reflected the reality of practice in the NHS. The fact is that its whole purpose was to do so, to the extent that any examination can, by recreating scenarios requiring clinical judgments that would arise in real life. The other part of the answer to this point is that, as noted above, the Respondent was not by this PCP seeking on its own and in isolation to achieve the legitimate aim. It was not and cannot be responsible for appointments made by Trusts and the regulation of daily practice by the GMC.

172.7. Finally, Mr Grundy suggested that the Respondent could have applied something like the Angoff technique to the Examination in the same way as it did

for the section 1 results. Although of course the Claimant was very close to achieving the pass mark, we do not know what would have been the outcome of any such adjustment or analysis. In any event, section 1 is a very different examination format to section 2. The pass mark for the latter is rigorously set at the calibration process, before the Examination is taken.

172.8. The Claimant failed by small margins, but it is agreed that a pass mark had to be set somewhere.

173. We can deal with PCP3 more briefly. Whilst it may not have been stated in the Respondent's briefing document for examiners, we find it wholly unsurprising that the standard applied in the Examination was that of a day 1 consultant in the UK. As some of the Respondent's witnesses pointed out, regulatory requirements and the nature of a day 1 consultant role will vary from country to country. Given that the Examination was (and is) part of the main route towards entry on to the Speciality Register of the GMC of the United Kingdom, we think it inarguable that the PCP served the legitimate aim – that is, it corresponded to a real need – and that it was rationally connected to it. The force of that point is not diminished by the possibility, mentioned by Mr Grundy only in closing and not put to the Respondent, that someone who passes the Examination was not bound to work in the UK. The whole point is that the vast majority will, for at least some period of time.

174. As to whether it was no more than necessary to achieving the aim, most if not all of what is said above in relation to PCP1 would also have been relevant in holding that PCP3 was also justified, including the multiple opportunities to pass the Examination and meet the required standard and the fact that it was open to the Claimant to apply for a training post, with the possibility of a reduced timescale for achieving the CCT. In addition:

174.1. As the Respondent submitted, it is difficult to see how the examiners could have assessed candidates to a different standard, or that it would have been appropriate to do so. We do not read Mr Woodward's statement as saying something different; he simply says that colleagues who come to work in the UK from another country can be more experienced than those in training posts.

174.2. We have noted above Mr Grundy's submission that the Examination should reflect life in the NHS. As far as PCP3 is concerned, that is the Respondent's point.

175. Finally, in relation to PCP4, Mr Ryan gave somewhat sketchy evidence that there had been research suggesting that the more attempts a candidate undertook, the more likely it was that they were demonstrating their ability to remember answers to questions and mere good fortune, rather than the merits of their clinical judgment. Sketchy though it was, that is indicative of a rational connection between the PCP and the aim of securing patient safety. As to whether it was no more than necessary to achieving that end, it would have been relevant to note that more than four attempts had been permitted in the past, but any such practice should be kept under review and it seems to us in any event a decisive point in the Respondent's favour that it appears to have been required of it by the GMC in order to take account of the research mentioned by Mr Ryan. In addressing the GMC's requirement, the Respondent identified that the maximum of three attempts was the "sweet spot", but gave some headroom beyond that by

adopting the limit of four, demonstrative of a proportionate approach. Of course that may have had an adverse impact on the Claimant – we cannot be sure given that we do not know whether he would have taken a further attempt and passed – but we can see the merit of drawing a line at some point, there is a marginal difference between the previous maximum of five attempts and the subsequent four and for the reasons we have given we think that on the balance of probabilities the Respondent would have discharged the burden upon it had it needed to do so.

176. In any event we would not have been persuaded that the Claimant brought his complaint about PCP4 within such further period after expiry of the usual time limit as was just and equitable. Whilst not of itself a determinative factor, the delay was in our judgment unexplained, particularly given that he was legally represented from early in 2020 and, as Mr Briggs pointed out, the case subsequently went through a preliminary hearing on time limits and exchanges of further particulars.

177. The delay was also long. Whilst that too is not of itself a determinative factor, the case law makes clear that it is an important one for consideration in most cases. Whilst as stated above we would have found that the Respondent discharged the burden of proof on justification in relation to PCP4, had it not done so we would have found that it was to a not immaterial degree prejudiced in putting its case. Mr Ryan was relying on an attempt to recall research from several years previously, in relation to a decision he does not appear to have been involved in making. The late introduction of the complaint, as Mr Briggs submitted, militated against the Respondent putting forward a primary witness of fact on this point.

178. For all of the above reasons, the complaints of indirect discrimination fail. We now turn to consider the complaint of direct discrimination.

Direct discrimination

Preliminary matters

179. The burden of proof rested with the Claimant to show that there were facts from which the Tribunal could conclude, in the absence of any other explanation, that he was discriminated against. In short, the question for us to determine was whether he was less favourably treated than a UK-national comparator, because of race. As noted above, Mr Briggs accepted on day 1 of the Hearing that “not being a UK national” was sufficient for these purposes.

180. Section 23 of the Act provides that there must be no material difference between the circumstances of the Claimant and his comparator, for the purposes of less favourable treatment. The comparator in this case was agreed to be hypothetical, namely a UK-national who provided the same answers as the Claimant in the November 2019 Examination. The Claimant led no evidence on the question of less favourable treatment compared with the hypothetical person, nor was it addressed in Mr Grundy’s written submissions. Indeed, the most that was put to the Respondent’s witnesses to this effect was that unconscious bias was at play and that they knew that the Claimant was a non-UK national (Mr Grundy used the phrase “overseas doctor”). Mr Grundy confirmed in oral submissions that his approach to this complaint was therefore to invite us to

focus on the "reason why" question, namely to find that the Claimant's race was a more than trivial reason why he failed the Examination in November 2019, and then, acknowledging that the question of less favourable treatment cannot be ignored, to draw the inexorable conclusion that he was less favourably treated once we have found that race played a part in it. There were a number of planks to the Claimant's case to that effect. We deal with each below in assessing whether he established a prima facie case, recognising of course the need to also step back and look at the overall picture when doing so.

181. One of the planks of the Claimant's case was the statistics we have considered above in dealing with indirect discrimination. We have not accepted his case that the statistical evidence demonstrates that non-UK nationals were put at a particular disadvantage compared to UK nationals in the arrangements for the Examination, and so any case for less favourable treatment (or indeed for race being a feature of the Respondent's decision-making) cannot rest on that basis.

182. Finally by way of preliminary comment, we acknowledge of course that discrimination can be unconscious as well as conscious. We accept too that not all of the examiners need to have discriminated against the Claimant in order for his complaint to succeed, though as we have noted the Claimant's evidence was that they all did. We should also note that the Claimant did not seek to advance any case based on any language barrier experienced by him during the Examination. Mr Grundy referred in his closing oral submissions to the impact of the Claimant's manner of speech being a feature in the marking of the Examination, but this was not put to the Respondent's witnesses and there was no evidence before us to suggest that there was a lack of comprehension by examiners of what the Claimant said, or by him of what he was asked.

Knowledge that the Claimant was not a UK national

183. The first plank of the Claimant's case to the effect that race was a factor in the conduct of some or all of the examiners was that they were aware that he was not a UK national when he appeared before them at their station. As the Respondent submitted, anyone of any ethnicity can be (or become) a UK national and first language is not necessarily a proxy for nationality. Whilst as Mr Grundy points out in his written submissions, Mr Ryan said that it was obvious when a candidate was not from the UK, the picture amongst the examiner-witnesses was mixed. Mr Kockelbergh said that a person's name may give some idea, Mr Crew said that on meeting a candidate it is possible to see it they are Asian or African (this most obviously relates to ethnicity, much less obviously to nationality), whilst Mr Mills said nationality was impossible to work out. Mr Shukla in his witness statement said that a candidate's ethnic origin was obvious on seeing them, but not their nationality, though in oral evidence he accepted that it would be obvious the Claimant originated from outside of the UK on the basis of his name, colour and accent. Assessing the evidence as a whole, whilst, as we have said, the picture is mixed, we are prepared to accept the Claimant's case, based on the evidence of Mr Ryan and Mr Shukla in particular, that his not being a UK-national would have been evident to the examiners, or at least to some of them.

Background matters

184. The Claimant also relies on a number of background matters which he says indicate that his treatment was because of race. In no particular order:

184.1. He says that the Respondent punished him because of his complaint about the second occasion on which he failed the Examination. That is not established on the facts, given that Mr Kockelbergh (one of the subjects of the complaints) was not informed of them, and in any event that would not be evidence of race discrimination, but of victimisation or something akin to it.

184.2. He also says that his experience as a urologist, his passing section 1 first time, his good references and the fact that he is working as a locum call into question the reason why he did not pass the Examination. Those are all facts of course, but none from which in our judgment we could conclude that he was discriminated against, for the simple reason that success in the Examination is very obviously not based on what a candidate has done (or is known by senior colleagues to have done) in practice to that point, still less on what they may do afterwards, but on their performance on the day – or, more particularly, the examiners' assessment of that performance. As Mr Briggs pointed out, passing section 1 and getting strong references are a separate part of the assessment of all candidates. On the same point, we note too that the Claimant's section 1 result was in percentage terms closely aligned to his result in the Examination, though as Mr Grundy says in his submissions, we do not know the details of the Claimant's marks. Furthermore, the Claimant's Examination result was not out of kilter with any of his section 1 result, his references or his broader experience: he was evidently not assessed to be a weak candidate – he failed the Examination by only two marks.

184.3. We do not think either that we could conclude that what the Claimant says about the body language of Mr Kockelbergh and Mr Woodward could contribute to a set of facts that would create a prima facie case of discrimination. This was not something that was emphasised by the Claimant either in his own evidence or when cross-examining the Respondent's witnesses – in fact questions about this matter came from the Tribunal – nor was it mentioned in Mr Grundy's submissions. Given the ambiguity of what the alleged body language might mean even if it did take place, the care which Mr Kockelbergh told us is required when a candidate was at or leaving a station, and particularly the fact that Mr Kockelbergh's marks were not out of kilter either with his co-examiner (Mr Crew) or any of the examiners generally and that Mr Woodward's were strongly aligned with Ms Gall's, we do not see that this is something that could lead to an adverse inference against the Respondent. We note too that body language is something which was covered in examiner induction training, or at least it was in October 2019 (see page 388).

184.4. As to the Claimant's colleague who passed on the third occasion, having obtained a UK passport after failing on the second occasion, even accepting that to be the case, we know nothing of that person's performance across his three examinations so that no safe conclusion can be drawn from it.

Examination marking generally

185. Whilst broadly speaking the Claimant did not seek to challenge the arrangements for the Examination and its assessment generally, there were two elements of those arrangements from which he invited us to draw adverse inferences.

186. The first was the mismatch between on the one hand the briefing for examiners at page 216 and the descriptors it provided for each mark, and on the other the marking sheets used by the examiners in November 2019, specifically the fact that the three domains on the marking sheet did not reflect the categories of assessment set out in the briefing document. That is of course an indisputable fact, but again we do not think it was a fact that could reasonably be said to contribute to an inference of discrimination. First, the document at page 216 was guidance. Secondly, there was no evidence that any of the examiner witnesses followed it in relation to other candidates but did not follow it for the Claimant. In fact, quite to the contrary, given their difficulty in recalling the Examination so many years later, all of them spoke very much about their general practice over years of examining and it is clear that all candidates at the Examination were assessed using the same template marking sheet. Thirdly, it was clear to us that it was the standard-setting or calibration exercise that was the key to producing a guide for examiners to follow at their stations when allocating marks to candidates. We did not see any examples of this guide (something we return to below) but all of the examiner witnesses said that this is what they used, adopting Mr Dickinson's grid in November 2019, which we unhesitatingly accept as unchallenged evidence.

187. The second was that the Synthesis domain was differently understood by each examiner. We do not think that in fact their understanding was as different as Mr Grundy made out. Mr Crew said it was about making safe and good decisions, Ms Gall that it was about creating a sensible management plan, Mr Shukla that the focus was what treatment option the candidate would recommend, and Mr Mills that it was about whether the application of knowledge was undertaken safely. It seems to us that whilst there were nuanced differences, there was thus a core commonality between them. We draw no adverse inference either from the fact that the Synthesis domain has now been discarded, as a number of features of the Examination appear to have evolved over the years, as Mr Briggs pointed out, including the use of the grid and the number of opportunities to pass. That kind of change is in fact what we would expect to see.

How the Claimant was marked

188. We will deal separately below with the actual marks the Claimant was given, but turn next to a number of features of the way in which the Claimant was marked at the Examination which he says should lead to inferences of race discrimination being drawn. None of the highlighted features have a racial element to them as such, and thus it is clear that the Claimant is inviting us to draw an inference from what he characterises as a series of failures by the Respondent to comply with its own guidance. We note in assessing these matters the case law, including **Bahl** and **Anya**, to the effect that not following guidance – which the Claimant implicitly equates to unreasonable behaviour – does not automatically mean that a prima facie case of discrimination is made

out, though it can lead to inferences being drawn where there is no explanation for it.

189. First, the narrative element of most of the Examination marking sheets that we were taken to contained no or limited reference to the marking descriptors set out in the briefing document, though with variation between different examiners. Notwithstanding that the briefing document was for guidance only and, as we have said, was effectively superseded by the calibrated question sheet (or, at the Examination, the grid), Mr Ryan said that the marking sheets should contain sufficient reference to the descriptors. Accordingly, whilst we suspect that the same approach was taken for all candidates, it is clear that the examiners did not follow the Respondent's own guidance consistently in this respect.

190. Secondly, the notes written by some examiners on the marking sheets did not relate to the domains against which they were recorded. We do not consider that to be of any significance. What the candidates, including the Claimant, said to the examiners at each station needed to be recorded and self-evidently what was said would not have fallen neatly within each consecutive domain as the answer unfolded. Whatever an examiner wrote had to be written on the marking sheet as the discussion progressed; it was the only sheet they had for recording the content of the interview. In an ideal world, the notes would have been in the relevant boxes and each domain separately assessed accordingly, but in a free-flowing examination, that would be very difficult if not impossible to achieve, as some of the Respondent's witnesses pointed out.

191. Thirdly, there were several instances where an examiner did not write down anything, or anything material, to justify a mark. This is related to the point about not referencing the marking descriptors from the briefing document. Giving a mark of 6 or 7 without comment was, we think, uncontroversial in practice, reflected to a degree in the briefing note which emphasised the importance of justifying a mark of 4 or 5 in particular. It seems to us that the Claimant himself takes a similar view in relation to instances where he was marked as a 7, in that we were not told of any challenge to Mr Kockelbergh's mark at page 714, where he gave a 7 without comment, Mr Shukla's mark of 7 at page 720 with the comment "Ok to good", the two marks of 7 at page 726 which simply recorded "Good" and "Sensible", and the mark of 7 at page 742 which recorded "Quotes evidence". It nevertheless remains that marks were given in some instances without adequate comment, which we accept was contrary to the Respondent's briefing note.

192. Fourthly, the Claimant was at pains to point out some inconsistency of marking in relation to his being prompted and making references to literature. We have detailed the relevant instances above.

193. Finally, Mr Grundy invited us to draw an adverse inference, to the effect that the marking was far too subjective and therefore susceptible to influence by considerations of race, from Mr Crew's evidence that he "gets a feeling" about a candidate, something similar being said by Mr Shukla. In our judgment, this submission invited us to place far too much weight on such comments. The word Mr Crew used was "impression". He was also careful to say that it was an early, not a first, impression and was clear that this related to the quality of the candidate's performance and confidence, so that he knew whether he could push them on to more complex answers after the opening questions. It must be

remembered that each scenario was discussed for only 10 minutes, which of necessity required a rapid judgment about a candidate's ability to deal with the scenarios. As Mr Shukla said, an initial judgment can change and the diversity training undertaken by each examiner addressed this point.

The Claimant's individual marks

194. The Claimant challenges a number of the marks he was given. We make clear that we are not competent to mark his answers, or seek to re-mark them. Our task is to determine whether any inference of discrimination should be drawn from the marks and the evidence given about them, either individually or taken overall. Our focus has been on those instances where there were differences between examiners on the same station, though we have also considered other challenges where the Claimant says a mark cannot be justified.

195. We begin with Mr Kockelbergh and Mr Crew, Mr Kockelbergh providing 37 marks overall and Mr Crew 35. The Claimant challenges their marks in a number of ways:

195.1. The first challenge is that Mr Kockelbergh gave him a mark of 6, not 7, in the Knowledge domain for scenario 1. We do not think that is a matter which can sensibly be said to give rise to, or contribute to, an inference of discrimination which requires an explanation in the sense envisaged by the burden of proof provisions. The simple fact that Mr Kockelbergh wrote down detailed notes cannot of itself suggest that a higher mark should have been given. Furthermore, there were two question marks in those notes and Mr Kockelbergh has given an explanation for his mark (that the Claimant did not know the histopathology of radiation cystitis and some of the treatments he referred to were historical) which we see no reason to go behind, not least when he gave the Claimant a higher overall score than his co-examiner and did not give him a mark lower than 6 in any domain.

195.2. Mr Kockelbergh now accepts that the Claimant did not have to be prompted about embolization and did know about hyperbaric oxygen but had to be prompted, so that there are some differences between what he wrote in his witness statement and the notes on the marking sheet. Mr Crew also accepted that his witness statement evidence was wholly inconsistent with what was recorded on his marking sheet, where the one comment he made was also wrong.

196. We turn next to Mr Mills and Mr Shukla, the former giving an overall score of 35 and the latter 37. Again, the Claimant levels a number of challenges at his scores:

196.1. Mr Mills said that his mark of 5 for Knowledge on scenario 1 was the result of the need for heavy prompting. We can unhesitatingly accept that as a clear explanation consonant with what he recorded at the time. It is true that Mr Shukla says he would have remembered had there been heavy prompting, but Mr Mills was the co-examiner and so would have been able to make a more detailed assessment of the Claimant's performance than Mr Shukla, whilst all Mr Shukla has recorded to justify his mark of 6 is the word "ok". It seems to us that it is therefore Mr Shukla's mark that requires an explanation, rather than Mr Mills', and of course it favoured the Claimant. Further, as we have noted above, Mr

Shukla ponders whether the difference was because of Mr Mills' greater expertise in the area, such that he would have looked for a fluent explanation of all of the basics. In view of all of the above, we have great difficulty with the criticisms levelled against what was after all only a difference of one mark and do not draw any adverse inference from it.

196.2. As for the difference of one mark in the Application domain for scenario 1, our analysis is the same. Again, we see no basis on which we could go behind Mr Mills' explanation that what the Claimant said was all basic urology and again, it seems to us that it is Mr Shukla's mark of 7 that would require explanation given that he wrote "ok – good", which at best seems to reflect a borderline score of 7. We note too that both examiners gave precisely the same marks for scenario 2, which suggests that they took a consistent approach to the candidate before them.

197. The Claimant says that there is no explanation for the differences in marks at stations 4 and 8 because none of the examiners for those stations gave evidence to explain their scores. We do not agree. At station 4, all marks from both examiners in all three domains were either 6 or 7, and the comments of the lower marker at page 726, though brief (he seems to have been the lead examiner), seem to us to entirely match the scores he gave – "fine", "prompting needed", "fine", "all fine" scored 6, and then "good" and "sensible" both resulted in a 7. Where the Claimant passed this station comfortably, even in the view of the examiner who gave the lower marks, we see no basis on which to draw an inference that he or she was influenced, consciously or otherwise, by considerations of race. As for station 8, the difference for Synthesis on scenario 1, which was the single difference between the markers, seems to be explained by the lower marker's reference to "limited progression with this scenario". With that explanation, and a one-mark difference overall, we see no basis on which to say that this should contribute towards an adverse inference being drawn.

198. In relation to station 3, Ms Gall and Mr Woodward, the Claimant's challenges to his scores are as follows:

198.1. He says first that Ms Gall failed to follow the briefing document by giving him a 5 for prompting, when the document suggests that minimal prompting should be marked as a 6.

198.2. Secondly, Ms Gall marked him down in two domains for the same reason on scenario 2.

199. We do not think there is great force in the Claimant's challenges on these points. First, like Mr Woodward, Ms Gall identified what was in the view of both examiners a major error in the Claimant's answer. It cannot be doubted that they heard him make reference to an invasive treatment, given that both wrote this down on their contemporaneous notes. Secondly, she and Mr Woodward were very consistent in their overall marks, differing by only one. That also seems to us to deal with the Claimant's point that he passed the bedwetting scenario in 2018 – and as we have said, it is performance at the Examination that counted. Thirdly, the Respondent's data indicates that Ms Gall is known as a slightly harsher marker overall than some of her colleagues. Fourthly, she was the only examiner witness in respect of whom we were taken to an analysis of how she scored BAME and White candidates respectively, which revealed no noticeable

difference. The scores from this station were not facts which could sensibly contribute to an inference of discrimination being drawn.

200. Finally there was a one-mark difference between Mr Hodgson and Mr Thomas. Although only one mark and although both examiners gave pass marks in every domain, we accept that Mr Thomas' statement does not provide much of an explanation for it.

Missing evidence

201. The final plank of the Claimant's case to the effect that adverse inferences should be drawn against the Respondent relates to missing documents and witnesses.

202. There are a number of documents that were mentioned during oral evidence that did not appear in the bundles:

202.1. The calibrated question sheets from either the Examination or some other diet. The additional document handed up by the Respondent towards the end of the oral evidence was a question sheet, but was not calibrated.

202.2. The grid produced by Mr Dickinson.

202.3. Possibly a urology-specific briefing note, something akin to the document at page 216, though we are doubtful that there was any such note as it was mentioned only by Mr Hodgson and then with considerable hesitation.

203. Mr Grundy's submission was that this prevented the Claimant from cross-examining the Respondent's witnesses about what the examiners had in front of them on the day of the Examination. We can see the force of that argument, particularly in relation to the grid, which seems to have been the guide for examiners on that occasion. We had no evidence as to what happened to these documents after the examiners handed them in. Whilst we were not taken to any evidence as to when, if ever, the Claimant first requested them, he did challenge his marks by way of his appeal immediately and brought this Claim relatively quickly thereafter. Whilst we accept Mr Briggs' submission that seeing something equivalent to the grid from subsequent diets would not have assisted the Tribunal, the Respondent's explanation as to why the 2019 grid was not available was incomplete, namely Mr Dickinson's comment that it is unlikely it was retained.

204. We are not persuaded that any adverse inference can or should be drawn from the Respondent's selection of which witnesses should appear before us to give oral evidence. We accept Mr Grundy's submission that there was no explanation for the absence of Messrs Thomas and Woodward, who it was plainly intended on exchange of statements and prior to the Hearing would give evidence, beyond Mr Briggs' general reference to diary issues (work commitments at the time of a junior doctors' strike) and "litigation decisions". We nevertheless accept, both in relation to them and in relation to the examiners from stations 4, 6, 7 and 8, the submission that their evidence would likely have been repetitious of their colleagues' and added little to the Tribunal's understanding of the case. Messrs Thomas and Woodward gave scores which were within one mark of their fellow-examiner and in Mr Woodward's case his

notes and statement give a clear explanation of his decisions. As to the examiners who gave no evidence at all, although it can be seen from the Case Management Summary of 12 January 2022 (page 102) that the Respondent indicated it would call up to twenty-five witnesses, most of the Claimant's challenges to his scores (well summarised in his further particulars at pages 72 to 75), certainly the more detailed ones and most of those directed at marks below the pass mark of 6, concerned the examiners who did give evidence. We were satisfied, given the Claimant's marks at each station, that the witnesses we saw were not unrepresentative of the whole picture. We add for completeness that we read nothing untoward in the decision not to call Mr Burgess as a witness given that he was not an examiner or in any other way involved in the Examination.

Overall picture at stage 1

205. All of this evidence having been analysed leaves the following as the basis of the Claimant's argument that he has established a prima facie case that race was a factor in the marks he was given at the Examination:

205.1. The fact that some examiners knew he was not a UK national.

205.2. The fact that many, it might be said most, examiners did not follow the Respondent's own guidance by not consistently referring to marking descriptors in giving their marks.

205.3. The fact that some marks were assigned without any or any adequate comment, contrary to the Respondent's briefing note.

205.4. The fact of some inconsistency between examiners on the question of how marks were influenced by levels of prompting and references to research literature.

205.5. The minor inconsistency between Mr Kockelbergh's witness evidence and his notes on the marking sheet and the complete inconsistency between the same for Mr Crew.

205.6. Mr Thomas' inability to explain a one-mark difference between Mr Hodgson and himself.

205.7. The absence of some documentation, specifically the grid championed by Mr Dickinson.

206. Given the complete absence of any actual evidence as to less favourable treatment of the Claimant, our assessment that the statistics we were taken to do not show any evidence of discrimination against non-UK nationals generally, our strong inclination from hearing the Respondents' witnesses to the view that the matters summarised at paragraphs 205.2 to 205.4 above would have applied to any candidate taking the Examination, and the minor deficiencies in the evidence of Mr Kockelbergh and Mr Thomas, we are not satisfied that what remains – namely the fact that some examiners knew the Claimant was not a UK national, the inconsistency in Mr Crew's evidence and the absence of some documentation – is sufficient for a reasonable tribunal to conclude that there was unlawful treatment and so pass the burden of proof to the Respondent. The

complaint fails on that basis. Even if it had been sufficient, indeed even if all of the matters on which the Claimant relied as we have analysed them above had been accepted as the basis of a prima facie case of discrimination, we would have been amply satisfied that the Respondent discharged the burden at stage 2, to which we now turn to conclude.

Stage 2

207. Although the Claimant levelled no criticism against the format, structure and overall approach of the Examination as such, it would have been highly relevant when considering the Respondent's case at stage 2 of the burden of proof to note a number of things which in our view contributed to a robust system of assessment:

207.1. First, as was agreed, the method of assessment (a "viva") is very common in medicine.

207.2. The process of standard setting or calibration for questions appears to have been detailed and rigorous process, taking place afresh at each diet (including for the Examination) over several hours and involving multiple people.

207.3. Performance at each station was marked independently by two examiners, something which is well-recognised to be helpful in mitigating against one person's impression being too influential.

207.4. The Claimant, in common with all candidates, was assessed at multiple stations so that in fact he and each candidate were seen and assessed by sixteen people. That further mitigates against individual bias playing a significant role in the outcome.

207.5. Each of the examiners in question had examined before, and so would have been assessed in their own performance (as examiners) both by assessors at their station and by the Respondent's analysis of how their marking compared to others and how it affected, if at all, people in different racial groups. That is something which would have helped to embed fair practice, particularly for someone who, like the Claimant, was a candidate on several occasions, meaning those who examined him had been repeatedly assessed by the time of the Examination.

207.6. The fact that the Respondent requires all examiners to undertaken equality and diversity training both at induction and before each examination diet, the latter shared between different consultants, is of course no guarantee against discrimination but is nevertheless to be commended as something which would generally tend towards eradicating it.

208. It is also in our view critical to a proper assessment of this Claim to emphasise that the Examination, which was an attempt to assess if someone was ready to become a consultant, did not lend itself to binary assessment of a candidate's performance. Whilst of course there was a structure (with some defects in practice that we have referred to) and a marking scheme, it is inescapable in an examination that seeks to recreate scenarios testing clinical judgment that the examiners' scores will be to a material extent matters of judgment themselves. No briefing or set of marking descriptors would be able to

take away the element of subjective judgment or make the practice of all examiners uniform. It is noteworthy that the Claimant accepted that two examiners could give different marks for the same answer, and the nature of the Examination is the reason.

209. Mr Grundy, understandably, advanced a case based on the need for a strict correlation between what the Respondent set down as guidance for examiners and their actual marking, but in our judgment that is an altogether too forensic approach to an examination of this nature. In our judgment:

209.1. It is inevitable that assessment in each of the three domains would overlap and unrealistic to expect the assessments to be hermetically sealed off from each other, particularly as the discussion of a scenario will not unfold in a neat and ordered fashion.

209.2. The question of the prompting required for a candidate is not a scientific judgment or tick box exercise based on adding up the number of prompts that were given. We accept that examiners can take different views from each other as to the significance of the prompting that was needed in any individual case. The same can legitimately be said to be the case in relation to references to literature, which we thought Mr Shukla explained well.

209.3. In the real world of clinical practice, which the Examination sought to recreate, consultants could legitimately take different views on a whole range of issues, for example the approach to advising a patient. It is artificial to expect anything like complete consistency.

209.4. In fact, we would expect some inconsistency in an Examination of this nature. A greater level of consistency, all the more so complete consistency, would justifiably arouse suspicion of collaboration between examiners either within or across stations.

209.5. In respect of what appeared on the Claimant's marking sheets, it seems to us inevitable that as a general rule the lead examiner would write much less than their colleague, focusing as they should on engaging with the candidate. More broadly, the time pressure in the Examination cannot be ignored in terms of how the ideal marking process contrasts with what often happened in practice. The requirement to mark a candidate immediately is a safeguard against misremembering what a candidate has said and/or comparing them with someone else.

209.6. Mr Grundy criticises the Respondent for only identifying and focusing on differences of two marks or more. Whilst we understand of course that from the Claimant's point of view, every mark is important, the Respondent's practice makes sense in the context of the judgment calls which as we have said the Examination inevitably required.

210. As to the Claimant's actual marks, we agree with the Respondent's assessment outlined by Mr Briggs that all the indications are he was scored on the basis of the examiners' assessment of what they heard and not, consciously or unconsciously, on the basis of his not being a UK national:

210.1. First, there is the fact that he was so close to passing the Examination. None of the examiners would have known what his overall mark was likely to be, so as to determine consciously or unconsciously what mark would be required to bring him to a point where he failed overall, there being no suggestion in the evidence to which we were taken that any of the examiners, either at one station or across different stations, knew what marks the other examiners were giving before they gave their own.

210.2. The same is true of the very strong broad consistency of marks given to the Claimant across the eight stations. There was no greater difference than a total of four marks on any single station – that was station 4, and in fact on all but that station (which represented the Claimant's highest score overall), there was no more than a total difference of two marks at any one station. That is a noticeable consistency.

210.3. Across all eight stations, there was no difference of more than one mark on any individual domain at any station.

210.4. The overall difference between lead and co-examiners if taken as two cohorts, was 12 marks, or around 4% of their respective share of the 576 marks required to pass (288 each).

210.5. To the above, one might add the fact that there was a substantial degree of similarity in the Claimant's overall marks across his four attempts at the section 2 examination and between that and his section 1 score.

211. We have noted and carefully considered the discrepancies between the Respondent's witness evidence and what is stated in the marking sheets. The only substantial discrepancy concerns Mr Crew, who accepts that the comment he put on the marking sheet was also wrong. Whilst we entirely accept that his mark sheet was an example of poor marking practice, we also note the following:

211.1. First, if the Claimant's case were to be accepted, we would have expected Mr Kockelbergh to have given him the lower mark, given that he was viewed by the Claimant as having marked him unfairly before and as someone who had given a pass mark to one of the Claimant's colleagues upon that colleague obtaining a UK passport. In fact, as we have noted, Mr Kockelbergh gave the Claimant a higher mark than was given by Mr Crew in both scenarios.

211.2. Secondly, in the domain where Mr Crew's comment was incorrect, he gave the Claimant the same mark as Mr Kockelbergh, whose explanation for his marks we have found sufficiently satisfactory (see below).

211.3. Thirdly, Mr Crew offered an explanation for the one domain where he scored the Claimant lower than a pass mark, referring to issues with histopathology and healing problems.

211.4. Fourthly, in addition to needing to allow scope for differences of judgment between examiners for the reasons we have already outlined, we agree with Mr Briggs that there has to be some acknowledgment on our part of the passage of time since the marks were given. As he submitted, there is a difference between justifying a mark at the time and three years or more later being able to explain it

by reference to one's notes, limited as they were in this instance given that Mr Crew was the lead examiner on that station.

212. We have noted the more minor inconsistency in Mr Kockelbergh's evidence but are satisfied that Mr Kockelbergh explained his marks, on the basis of the prompt that was required in respect of the dilated ureter, the Claimant not getting through the full scenario and his concerns about the mention of the ureteroscopy. He must be permitted sufficient entitlement to clinical judgment.

213. In summary, though a small number of the marks given to the Claimant are not explained in detail, we are satisfied that seen in the context of what is set out above – in particular the nature of the Examination, the multiple safeguards against discrimination that the Respondent put in place, the time-pressured context in which the marking occurred, and the significantly consistent (but, tellingly, not uniform) assessment of the Claimant overall – and combined with the explanations that have been given, the explanations offered by the Respondent would have been more than sufficient to discharge the burden of proof upon it had it been required to do so. Some (or even all) of the examiners knowing the Claimant was not a UK national and the absence from the disclosed documentation of the grid used at the November 2019 Examination would not have led us to conclude otherwise.

Conclusion

214. We cannot but have sympathy for the Claimant given that he came so close to passing the Examination, and we have identified a number of respects in which the Respondent's practice at the Examination fell short of what would have been ideal, which we hope it will note and address. We are satisfied however, for the reasons we have given, that the Respondent did not discriminate against the Claimant, directly or indirectly. His complaints are therefore dismissed. The Remedy Hearing provisionally scheduled for 4, 5 and 6 December is cancelled and the Case Management Orders made in relation to the same rescinded.

This was in part a remote hearing (see paragraph 20). There was no objection to the case being heard, in part, remotely. The form of remote hearing was V - video.

Employment Judge Faulkner
Date: 9 August 2023