PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Association of professional identity, gender, team understanding, anxiety and workplace learning alignment with burnout in junior doctors: a longitudinal cohort study
AUTHORS	Monrouxe, Lynn V; Bullock, Alison; Tseng, Hsu-Min; Wells, Stephanie

VERSION 1 – REVIEW

REVIEWER	Nancy Sturman University of Queensland, Australia
REVIEW RETURNED	03-Jun-2017

GENERAL COMMENTS	Many thanks for this work, and the opportunity to review the paper. The collection of longitudinal data, the use of validated measures, and the focus on understanding an organisational factor as well as individual factors, are strengths of this study of an important issue The findings of increased burnout with a higher general predilection for anxiety, and reduced strength of professional identity (PI), are important and are supported by previous literature. As the authors suggest, further qualitative work is indicated to understand how to best address these issues during undergraduate medical education (and arguably consider them in medical programme selection). The finding that PI did not change over time during the study was interesting and somewhat surprising, suggesting that much of this formation occurs during medical school, which is counter-intuitive and perhaps could be compared to other literature . I must admit to some confusion about the key findings in relation to aligned versus non-aligned assistantships, which seem somewhat contradictory even after careful reading of the paper. For example, the abstract contains the statements "PB and WRB increased over time, but no effects of alignment" and "participants experiencing MAs reported higher PB and WRB over time". Again, "The findings around PI, teamwork and burnout suggests that assistantship alignment appears to provide no absolute beneficial effect", whereas the linear mixed-effect model "suggests a slightly delayed beneficial effect". (If MA intern burnout scores do increase more steeply over the intern year, I'm not sure how this can be explained by the alignment of the assistantship; it seems more intuitive to expect any effect of this alignment to dissipate with time. I am also unsure how the authors explain a delay in the benefit) The authors may be able to clarify this reader's confusion by some editing of the main text. The authors acknowledge that participants undertaking MA assistantships may have differed in other ways, including the degree of social disruption during the intern transition but I felt they tended

	to discount the importance of these differences, and that there may be other factors related to students who are not placed in their "home country" workplaces (or who undertook their medical school study outside their "home country") which could be posited. So although I agree that it is interesting to study this natural experiment, and there may indeed be a case for more protracted shadowing or other measures to assist junior doctors who are displaced from their country of undergraduate study and student workplaces, I am not sure that "a strong relationship" has been identified as stated by the authors, or that strong conclusions can be drawn from the data about the alignment of the assistantship per se. I agree however that there is a case for undertaking further studies internationally on "misaligned" junior doctors to understand any increased risks of burnout, and identify means of mitigating these. Please check age of participants in abstract, which implies some students were 17 yrs at T1. I wonder if there may be some inaccuracies in the Participants section of the Abstract, which is confusing to read, and suggests that a much smaller proportion of Aligned than Misaligned students consented to participate, and that more Misaligned students consented than were contacted. Interesting and well written background. The authors mention "group identification" in line 48 of page 4, and I wasn't clear how this related to PI (if at all) and identity threat (which would benefit from some explanation), nor whether the reference was to inter-professional or mono-professional healthcare groups. I would be interested to know why the authors chose the TUS in particular to explore team-work, given its very general nature. It seems to have little relevance to teamwork in workplaces or healthcare teams. Please correct typos in the scale's items, including item 4 which is unintelligible as it stands; please note that it is termed "Teamwork Scale" in the supplementary files and TUS in the text. In Results, please check statement that 46% of participants experienced the AA (see my comments on the abstract). I would be interested to hear the authors' views on the finding that males reported higher PRB levels, which appears to be a new finding. Could there be a gender-specific social desirability explanation here? Please correct Title of Figure 1. Please note that when printed in black and white the Figure is not as clear to the reader as perhaps anticipated. Overall this is an interesting, well conceived and well written study which warrants publication. Thank you for the opportunity to review it.

REVIEWER	Bryan Burford Newcastle University, UK
REVIEW RETURNED	16-Jun-2017

GENERAL COMMENTS

Overall
This is an excellent paper addressing an important and topical area. However, I did find the results section rather brief, and think the

	transition for individuals with a predilection for anxiety. I think the authors wish, in the first sentence, to say something like: ³2XU ILQGLQJV VXJJHVW PHGLFDO VFKRR year work place alignment before graduation, or an extended VKDGRZLQJ SHULRG LPPHGLDWHO\ DIWHU VHQWHQFH , ZRXOG XVH µSUHGLVSRVLQ¶ µ3UHGLOHFWLRQ¶ VXJJHVWV D SRVLWLYH 3DJH ³0HQWDO KHDOWK SUREOHPV L XQGHUJUDGXDWH OHYHO KDYH SRWHQWL , WKLQN WKH DXWKRUV PHDQ µLPSOLFDW 8) 4:40 The professional socialisation of medical students has been of focus recently. µ:KDW GRHV µRI IRFXV¶ PHDQ"¶ 9) 4:43 develop a strong, sense of themselves as doctors. Stray comma. ³VDUHWDWHG LQFLGHQFHV´ , EHOL µLQFLGHQW¶ ³$GGLWLRQDOO\ WKHUH LV D JHQ UHVSRQGHQW QXPEHUV DV ORZ DV Q This would be better expressed as: ³$GGLWLRQDOO\ UHVSRQGHQW QXPEHUV n=43 at WLPH ´ 12) 5:42 In some qualitative studies, participant numbers are also relatively low (i.e. n=15 and n=20). I would say that n=20 is low for a quantitative study, but may be excellent for a qualitative study, depending on the approach taken. 13 µDQG LQ GLIIHUHQW FRQILJXUDWL ³\RXU VWRU\´ µPLVDOLJQHG DVVLVWDQWVKLSV¶ µPLVDOLJQHG¶ VXJJHVWV WKHUH LV D SDVVLYH RQH , ZRXOG UHF-RPPHQG accurate descriptor. µD OL-QPNDZDUGH¶ 3DJH µ3URFHGXUH¶ The second and third paragraphs should be re-written into one paragraph to eliminate duplication. 19) 7:7 Font size change µ$ WRWDO¶ 21) 8:7 An total of n=281 (56% response rate aggregated over time)

	participated in this study, comprising: 68% (n=183) females and 73% (n=206) aged 17-25. 22) Here and elsewhere, the use of the colon seems unnecessary. This sentence works better as ³7 Kh response rate aggregated over time was 56% (n=281), of whom 68% (n=183) were female and 73% (n=206) were aged 17-´ ³7KH SHUFHQWDJH RI IHPDOH UHV 7KXV WKH VL]H RI WKH LQWHUFHS ³LV OLNHO to the Research characteristics of EXUQRXW´ , GRQ¶W IROORZ ±WKVHQ¶WDR JLLW GRXWKLV LQG ³7KH $$ SDUWLFLSDQW JURXS reported PB scores increased and an a YHUDJH RI SHU WLP What does this mean? 27) The authors may well be so familiar with their acronyms that they can decode them immediately. However, the reader is not in this position. Spelling out the acronyms will add slightly to the word length but very greatly to the clarity. 28) The authors report an effect of aligned versus non-aligned assistantships on the properties they measure. However, this is not under the control of the medical schools, Foundation programmes, or even entirely under the control of the graduates ±it results from the interaction between student preferences and their EPM and SJT scores in the Selection for Foundation process. Therefore, there is QRWKLQJ WKH V\VWHP FDQ FXUUHQWO\ G assistantships, and the observations made in the paper are interesting but currently not useful. The authors propose an extended shadowing period post-graduation (presumably between the June exam and the beginning of August) but (a) there is no evidence that this will work ±it is purely speculative and (b) there are challenges to the practicality of this approach, given that it will require graduates to move to where they will be working in F1, before they are paid a salary. 29) In order to indicate the generalisability of the findings, it would be valuable if the authors indicated for the UK generally, how many graduates work locally to their medical school. This number varies greatly (for instance, London has a low match), and the data is available from UKFPO.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1 Comments

Comment: The finding that PI did not change over time during the study was interesting and somewhat surprising, suggesting that much of this formation occurs during medical school, which is counter-intuitive and perhaps could be compared to other literature

Response: We agree that the idea that professional identity does not change in the study sample over the one-year period that we conducted the research is somewhat counterintuitive. One would expect that professional identity would increase over the time period in which an individual actually enters the profession.

We do not believe, however, that it would be fruitful to compare this finding to other literature. Mainly because of the differing ways in which the literature measures professional identity: the concept of professional identity is extremely popular within medical education and, as has been argued elsewhere, researchers tend to develop their own ideas of what professional identity comprises and µPHDVXUHV¶ RI LGHQWLW\ YDU\ JUHDWO\ DFURVV VWXGLHV. It is therefore difficult to compare across studies H[WUHPHO\ GL

What we can do is try to explain our findings by examining specific professional identity scale we have used and what we have found.

When we consider our scale (Appendix A), we notice that the items on the scale mainly relate to general professional identities around an individuals' alignment with the profession, the importance of the profession to that individual in terms of its value to them personally and their feelings of sharing similar characteristics with other members of the profession. When we consider our findings, what we DUH VD\LQJ LV WKDW SDUWLFLSDQWV¶ WKRXJKWV DURXQG WKHVH LV SHULRG VWXGLHG :H GR QRW VD\ KRZHYHU WKDW WKHLU LGHQWLW\ GRHV QRW FKDQJH RYHU WKDW WLPH UDWKHU ZH FRPPHQW WKDW ³WKLV VKDUHG LGHQWLW\ PLJKW KDYH E VWXGHQWV WUDQVLWLRQ LQWR SUDFWLFH´)XUWKHUPRUH ZH GR IL feeling of identity ±as we have demonstrated in our linear-effect model. We have now amended this FRPPHQW WR LQFOXGH WKH FDYHDW ³IRU VRPH WKLV VKDUHG LGHQW

We have also included a comment in the discussion about those who have not yet formed a strong identity ±alongside a recognition that we did not measure identity prior to the assistantship (as T1 was timed in the middle of the period):
³)RU WKRVH SDUWLFLSDQWV ZKRVH VKDUHG LGHQWLW\ ZDV LQVXIILF appears that such identities, likewise, are nether diminished or enhanced over time leaving them YXOQHUDEOH WR H[SHULHQFLQJ EXUQRXW +RZHYHU DV ZH GLG QRW professional identities prior to the assistantship programme (T1 being in the middle of it), we cannot EH IXOO\ VXUH WKDW WKH HDUO\ ZHHNV GLG QRW KDYH DQ LPSDFW R

Comment: I must admit to some confusion about the key findings in relation to aligned versus non-aligned assistantships, which seem somewhat contradictory even after careful reading of the paper. For example, the abstract contains the statements "PB and WRB increased over time, but no effects of alignment" and "participants experiencing MAs reported higher PB and WRB over time".

Response: We have now clarified this in both the abstract and the methods section:

ABSTRACT:
³5HVXOWV $QDO\VHV RI LQGLYLGXDO JURZWK FXUYHV GHPRQVWUDWHG VHOI patient-related burnout were stable, while team understanding, personal and work-related burnout increased, all irrespective of alignment. Three linear mixed-effect models (personal, patient-related and work-related burnout as outcome measures; age and gender as confounding variables) found that males self-reported significantly lower personal, but higher patient-related burnout, than females. Age and team understanding had no effect. Anxiety was significantly positively related and

professional identity was significantly negatively related to burnout. Participants experiencing non-aligned assistantships reported higher personal and work- U H O D W H G E X U Q R X W R Y H U W L P H ´

METHODS:
³ $ Q D O \ V L V F R P S U L V H G W Z R V W D J H V 6 W D J H D Q D O \ V L V Z D V X Q G H U W
self-reported professional identity, anxiety, team understanding and burnout levels changed over the time of our study as final year students transitioned into their clinical practice. Here, the individual scales were examined (for the Copenhagen Burnout Inventory the three sub-scales of burnout were examined separately: personal, patient-related and work-related). The significance of the main effects (alignment group and time) and the interaction effects for aligned/misaligned groups over time were assessed with the Wald chi-square test using GEE estimation.

Stage 2 analyses was undertaken to answer RQ2 by examining the associations between assistantship alignment, self-reported professional identity, anxiety and team understanding on E X U Q R X W V F R U H V R Y H U W L P H ³

Comment: Again, "The findings around PI, teamwork and burnout suggests that assistantship alignment appears to provide no absolute beneficial effect", whereas the linear mixed-effect model "suggests a slightly delayed beneficial effect".

If MA intern burnout scores do increase more steeply over the intern year, I'm not sure how this can be explained by the alignment of the assistantship; it seems more intuitive to expect any effect of this alignment to dissipate with time. I am also unsure how the authors explain a delay in the benefit

The authors may be able to clarify this reader's confusion by some editing of the main text.

5 H V S R Q V H : H S R O L W H O \ G L V D J U H H Z L W K W K H D V V H U W L R Q W K D W ³ L W
this ali J Q P H Q W W R G L V V L S D W H Z L W K W L P H ´ , Q G H H G W K D W W K H D O L J Q H
work-related burnout on their first transition to being a junior doctor is not unexpected. Indeed, the sharp rise in responsibility as perceived by new graduates is well documented and it is highly possible that students undertaking an aligned assistantship might feel more protected prior to their first post beginning (we have found this in our qualitative data previously published). So, there might be an initial µ V K R F N ¶ D W I L U V W I R O O R Z H G E \ D S H U L R G R I J U H D W H U V W D E L O L
study findings (on page 10).

Comment: The authors acknowledge that participants undertaking MA assistantships may have differed in other ways, including the degree of social disruption during the intern transition but I felt they tended to discount the importance of these differences, and that there may be other factors related to students who are not placed in their "home country" workplaces (or who undertook their medical school study outside their "home country") which could be posited.

So although I agree that it is interesting to study this natural experiment, and there may indeed be a case for more protracted shadowing or other measures to assist junior doctors who are displaced from their country of undergraduate study and student workplaces, I am not sure that "a strong relationship" has been identified as stated by the authors, or that strong conclusions can be drawn from the data about the alignment of the assistantship per se.

Response: We have clarified the situation in our methods section. We do believe that that we have I R X Q G V W U R Q J U H O D W L R Q V K L S V D Q G W K D W W K H µ Q D W X U D O H [S H U L P
confounding variable ±this is because although the aligned assistantships are for only those studying L Q : D O H V D Q G Z K R V H I L U V W S R V W L V L Q : D O H V W K H \ D U H Q R W Q H F H

some of the graduates from outwith Wales are moving back to their home country. As such both groups might experience a disruption to their social situation. We have amended as follows:

³)RU XQGHUJUDGXDWHV ZKR ZLOO FRPPHQFH WKHLU SRVWJUDGXDWH aligned in location and specialty with their first post-graduate role as a junior doctor (intern) post. This first post is not necessarily in the same location within the country. For example, the graduate could be starting their new junior doctor post in a remote northern Wales setting despite them undertaking the majority of their studies in a southern city setting. Welsh undergraduates who will commence their first post in another UK region and undergraduates from other UK medical schools who will commence their first post within Wales undertake non-aligned assistantships. Some of these graduates might be returning to their home country, Wales. This context provides us with a natural experiment situation, with both groups of participants likely experiencing some form of disruption to their social sett L Q J ´

Comment: Please check age of participants in abstract, which implies some students were 17 yrs at T1.

I wonder if there may be some inaccuracies in the Participants section of the Abstract, which is confusing to read, and suggests that a much smaller proportion of Aligned than Misaligned students consented to participate, and that more Misaligned students consented than were contacted

Response: We thank the reviewer for noticing this and for querying the ages of participants. The ages are at time of entry, rather than graduation. We have now amended our abstract as follows:

³$OO DOLJQHG DVVLVWDQWVKLS VWXGHQWV (n=319) were contacted. n=281 (56%) responded: 68% (n=183) females, 73% (n=206) 22-30yrs, 46% aligned Q ´

Comment: Interesting and well written background. The authors mention "group identification" in line 48 of page 4, and I wasn't clear how this related to PI (if at all) and identity threat (which would benefit from some explanation), nor whether the reference was to inter-professional or mono-professional healthcare groups.

Response: Thank you for this query. In terms of the first query regarding group identification and being unsure as to how it relates to professional identity ±in our original introduction we state:

³7KXV DQ HVVHQWLDO SDUW RI EHFRPLQJ D GRFWRU FRPSULVHV VWX being part of the medical profession:29 30 32 33 that is, developing an embodied professional LGHQWLW\ ZLWKLQ WKH DGGHG WKH SKUDVH ³WR DQG WKHQ SHUVLVW WR further.

As for the issue of identity threat ±we have clarified this with further details as follows:

³*LYHQ WKDW WKH JURXS LV SDUW RI RXU LGHQWLW\ we can include feeling that the group we identify with is threatened in some way or feeling alienated from the group GLVUXSWLQJ RXU VHQVH RI EHORQJLQJ ´

Comment: I would be interested to know why the authors chose the TUS in particular to explore team-work, given its very general nature. It seems to have little relevance to teamwork in workplaces or healthcare teams. Please correct typos in the scale's items, including item 4 which is unintelligible

as it stands; please note that it is termed "Teamwork Scale" in the supplementary files and TUS in the text.

Response: We chose this scale as it relates to specific understandings of team-working and was developed alongside the professional identity scale we have used. We have now addressed the labeling issue and refer to it throughout the manuscript at the appropriate junctions. We have also amended the wording in this section and hope make the sentence more intelligible.

Comment: In Results, please check statement that 46% of participants experienced the AA (see my comments on the abstract).

Response: This figure is correct as this percentage is from the total number of our participants rather than the total number of participants experiencing aligned assistantships that existed in 2015.

Comment: I would be interested to hear the authors' views on the finding that males reported higher PRB levels, which appears to be a new finding. Could there be a gender-specific social desirability explanation here?

Response: We thank the reviewer for highlighting this aspect and have now added in our consideration of a social desirability explanation to the discussion. One explanation for these gender-specific findings might be that respondents are acting according to their gendered expectations. Thus males might downplay distress due to a social expectation that they should appear tough and resilient.76 Thus, in our study we find that males report greater patient-related burnout but females reporting greater personal burnout. In terms of the social expectations explanation, it could be that males are downplaying their personal burnout (subconsciously acting out a tough male role) but up playing patient-related burnout (i.e. subconsciously shifting from self to others). Conversely, for females (i.e. subconsciously due to societal expectations around their nurturing role). However, we remain agnostic on this interpretation, as it is impossible to really understand what is happening. Importantly, if it is that males are downplaying their experiences of personal burnout, they might be at greater risk of experiencing negative emotional wellbeing. As such we should not necessarily accept the findings in terms of J

Comment: Please correct Title of Figure 1. Please note that when printed in black and white the Figure is not as clear to the reader as perhaps anticipated.

Response: This has been corrected and the colours amended.

Reviewer 2 Comments

Comment: This is an excellent paper addressing an important and topical area. However, I did find the results section rather brief, and think the details of the regression analysis could be more clearly spelled out in both method and results.

Response: See further comments below

Comment: A clear and comprehensive review. One slight quibble is that the implicit hypothesis that professional identity may vary with assistantship alignment is not clearly justified. As it is not stated as an explicit hypothesis that may not be too much of a problem, but if there is a theoretical rationale, it would be helpful to have it clearly presented.

5HVSRQVH :H SUHVHQW WKLV WKHRUHWLFDO UDWLRQDO LQ WKH SUH
individual to institutional factors f RU EXUQRXW¶ DQG µFXUULFXOD LQWHUYHQWLRQ¶
for contextual factors affecting burnout and we also discuss other research that has attempted to examine the effect of assistantships on workplace stress. We believe that this information is sufficient to inform our hypothesis.

&RPPHQW , WKLQN WKH WHUP µLGHQWLW\ ±what paradigm? Re. IS FRXOG GR
UHDGV DV LI LW UHODWHV WR WKH WHDPZRUNLQJ HOHPHQW EXW WR
uVH RI µLGHQWLW\ WKUHDW¶ LQ WKH VRFLDO LGHQWLW\ OLWHUDWXU
may be the term in being used in a different sense. This para could also serve to better seed the LQFOXVLRQ RI WKH µ7HDP 8QGHUVWDQGLQJ¶ VFDOH

Response: We thank the reviewer for this comment. Please see our responses to the similar query DERYH DQG RXU GHYHORSPHQW RI WKH µLGHQWLW\ WKUHDW¶ FRQFHS

&RPPHQW , DOVR ZDVQ¶W VXUH ZK\ \RX KDG QR K\SRWKHVHV 54V UH
your outcome variables. While this would have added to or possibly changed the analysis, there is literature regarding the relationship between identity and wellbeing which may enhance the reach of your work.

Response: We have now clarified this section (see further information above) and have clearly linked the research questions to the two phases of the analyses.

Comment: Given the reference in the introduction to the effect of clinical area on wellbeing, was location or specialty recorded, and could that have been included in the lmer modelling? (I realise F1s are likely to be almost exclusively in acute hospitals, but an effect of site or specialty may still be possible).

Response: This is an interesting question ±but not one that we are able to answer at present. Further research might unpack this issue.

Comment: Under the analysis section, no details are given of model/variable selection ±I infer that no comparison of subsetted models was undertaken (to test for example the contribution of the interaction effects), which as I understand the analysis is acceptable, but should probably be stated clearly. I found myself wondering about interactions between PI, anxiety and time, and it would be nice to know if these were considered or excluded.

Response: We did include the interaction effect of Group and Time in our model (see the formula in P7) and reported the interaction effect in Table 3 (in the column titled Group * Time of P14). However, we did not do the interaction effect for subsetted models between time and other independent variables (i.e. PI or anxiety). We did not do this for two reasons: 1) the personal characteristics in PI or anxiety were taken as confounding variables for the effect of time and group which we were most interested in; 2) it would become extremely complicated due to the number of variables we have.

7R FODULI\ WKLV LQ RXU PHWKRG ZH DGGHG D VHQWHQFH LQ 3 ³$O\ interaction effects between these factors and time, we do not include these interaction terms in our PRGHO IRU VLPSOLFLW\ SXUSRVH ´

Comment: I think a plot of the continuous predictors and outcome variables would be helpful to clarify the significant relationships between those variables.

Response: We have now created several plots concerning the relationship between the continuous predictors and outcome variables. The plot of linear relationships among these variables is included in the newly created Appendix B. Followed by these graphs is another type of plot for the fixed effect in the mixed effect model which shows all the significant levels of all predictors. We include this in Appendix B too. As we are unsure if these really are helpful for the majority of readers, we suggest that if they are included in the final published version of the manuscript that they are retained within an Appendix.

NOTE: at the moment we do not refer to this Appendix in the manuscript and await the editorial decision on this. In any case, due to the open nature of this review, interested parties can access the plots if they follow the peer review documents.

Comment: It should be clearly stated that time is treated as a continuous variable in the model ±this confused me in the Results and came as a surprise in the discussion! While on further reading it seems it is appropriate in such a design, I think a clear statement and reference to that effect would be helpful to those like me who know something of lmer, but may not be familiar with every usage. On a similar note, inclusion of R code for the final model would be helpful for interpretation.

Response: We thank the reviewer for their comments and we now specify that we have treated time as a continuous variable when we discuss the model on page 7.

The R code is as follows (again, we are unsure if this is of general use and so await the editorial decision on if and how to publish this):

```
Burnout.Long.complete <- read.csv("~/Documents/University life/Research
Work/Correspondence/Lynn/Anxiety/R for Anxiety/Burnout Long complete.csv")
# write to a container for easily operation afterwards
mydata <-Burnout.Long.complete
# install necessary library for this analysis
library("lattice", lib.loc="/Library/Frameworks/R.framework/Versions/3.4/Resources/library")
library("lme4", lib.loc="/Library/Frameworks/R.framework/Versions/3.4/Resources/library")
# making factor variables incl. Group, Gender
group.f <- factor(mydata$Group, levels = c(1,2), labels = c("Yes","No"))
gender.f <- factor(mydata$Gender, levels = c(1,2), labels = c("Female", "Male"))
# descriptive statistics
# 3-Way Frequency Table
mytable <- table(gender.f, group.f, phase.f)
ftable(mytable)
# Summary Statistics by Group, Using Psych
library("psych", lib.loc="/Library/Frameworks/R.framework/Versions/3.3/Resources/library")
describeBy(mydata,list(phase.f,group.f),mat=TRUE,digits=2)
# make a new set of data (selecting cases)
newgr2data <- mydata[ which(mydata$numphase>1), ]
group2.f <- factor(newgr2data$Group, levels = c(1,2), labels = c("Yes","No"))
gender2.f <- factor(newgr2data$Gender, levels = c(1,2), labels = c("Female", "Male"))
```

# examine model fit by comparing Linear vs Quandric vs Cubic Need to verify
# Mixed Effect Model for burnout
pbsallgr2 <-lmer(PBS ~ group2.f * Phase + Age + gender2.f + HARS + PI + TW+ (1+Phase|ID), newgr2data, REML=FALSE)
wrballgr2 <-lmer(WRBS ~ group2.f * Phase + Age + gender2.f + HARS + PI + TW+ (1+Phase|ID), newgr2data, REML=FALSE)
prballgr2 <-lmer(PRBS ~ group2.f * Phase + Age + gender2.f + HARS + PI + TW+ (1+Phase|ID), newgr2data, REML=FALSE)

Comment: , ¶P QRW FOHDU RQ WKH SUHVHQWDWLRQ RI WKH Estimated Marginal Means above and why the effect of WL to be a categorical variable and wondered why contrasts were missing. As the random slope is not significant, I wonder if it may be clearer to include Time as a discrete fixed effect and present the contrasts rather Table 2, so simplifying the section?

Response: See above

Comment: RESULTS: The first para on respondent characteristics is a bit confusing. I would find a clear statement of response rate at each time point more informative.

Response: We now begin our results section as follows:

³7KH UHVSRQVH UDWHV IRU WKH IRXU TXHVWLRQQDLUHV ZHUH DV IRO
7 Q DQG 7 Q ´

Comment: Details of drop-out within each time point would be informative ±while not definitive, Bristol Online Surveys gives numbers of those who do not complete a questionnaire, and given the length of your form, it would be reassuring to know how few or many may have been deterred from completion by its length.

Response: As you correctly say, BOS will give out information regarding incomplete surveys. However, as it is designed to maintain respondent anonymity it will not automatically capture any identifying information about the respondents, we have to set up a respondent list apriori ±it can then tell us if a respondent started but did not complete the questionnaire. As we did not have access to SRWHQWLDO SDUWLFLSDQWV¶ HPDLO DGGUHVVHV we were unable to send reminders to WKH VWXG\ we did not set up the respondent list and so this information is not available to us.

In terms of individuals not completing all four of the time points, we do not think that survey length had anything to do with it. We had some problems with institutional email addresses expiring- this may have played some part in the drop off rate. Also, if people did not complete at T1, then we did not have a direct way to contact them at T2 onwards if they had left Wales. We will have lost people to follow up in that way.

Comment: , ¶P QRW VXUH ZK\ WKRVH ZKR FRPSOHWHG RQO\ RQH TXHVWLR effects modelling is not troubled by missing data, my understanding is there is no need to remove data like this. (My understanding may of course be wrong).

Response: Indeed, as the reviewer stated the advantage of mixed effect modelling is that missing data are ignorable. However, because we use time as a continuous variable, and the model effectively fills in the missing data, we wanted to be more conservative and provide the model with at least 2 time points. We believe that this is a more robust way of handling the missing responses as such we are more confident about missing responses can be explained by covariates in the model or by the available responses from a given subjects.

Comment: Finally, the random intercept for participants indicates to me individual difference rather than anything to do with the measurement characteristics of the scales (which to me implies something about their validity/reliability).

Response: We have amended the description to clarify this as individual differences for burnout measures.

Comment: The discussion is clear and pertinent. I think it could be a bit more concise (to the tune of quarter of a page or so) in recapping the results, and in spelling out clearly what the regression means ±ie regardless of time, there is an effect of anxiety, PI and teamwork, but the effect of alignment can vary over time, etc.

Response: We have now developed the Discussion to ensure it clearly sets out our study findings. However, due to the additions required by all three reviewers we have failed in terms of reducing the Discussion by one quarter of a page. We believe that clarity is more important and have been careful not to be overly wordy when ad G U H V V L Q J U H Y L H Z H U V ¶ F R P P H Q W V

Comment: The linkage to the wider context and literature seems sound and comprehensive, to my reading.

Response: Thank you.

Reviewer 3 Comments

Comment: This reports a study of an important area, and is certainly worth reporting. However, the Abstract in particular is confusingly written, and the title does not reflect the study. This is the major reason I have selected 'major revisions' as my recommendation. The use of similar acronyms throughout is particularly confusing and, at the cost of a few extra words, I would recommend writing them out in full each time, which would add greatly to the clarity of the paper - most readers will not be as familiar with them as the authors.
The recommendation that for graduates moving to a new area would benefit from an extended assistantship after graduation is plausible but speculative - no evidence in the paper supports this, and this should be made clear. Such a programme poses significant practical challenges, and this should be made clear in the paper. I believe the term 'non-aligned' is better than 'misaligned', since the latter suggests an active mismatch. A detailed list (repeating some of the above) is provided below, including typos.

Response: See below for our response to this as all points are repeated in further detail.

Comment: 1) 7 K H W L W O H L V ³ $ V V R F L D W L R Q R I S U R I H V V L R Q D O L G H Q W L W \ Z R U N S O D F H O H D U Q L Q J Z L W K E X U Q R X W L Q M X Q L R U G R F W R U V D O R Q J designed ar R X Q G D F R P S D U L V R Q D O L J Q H G Y H U V X V µ P L V D O L J Q H G ¶ D V V L V least make it into the title?

Response: Yes, we fully agree with this comment but have struggled to include the concept whilst minimizing the word-count. We have now D G G H G W K H W H U P µ D O L J Q P H Q W ¶ W K H W L W O H information as to the type of workplace learning we refer to.

Comment: 2) Design: A longitudinal one-year cohort design. An online questionnaire included: Professional Identity Scale (PI), Team Understanding Scale (TU), modified Hamilton Anxiety Rating

Scale (HAM-A) and modified Copenhagen Burnout Inventory (CBI). Data was collected on four occasions: (T1) prior to graduation

3) PB, WRB and PRB are not mentioned in the design, and HAM-A and CBI are not mentioned in the Results. They appear, as these are key part of the Results.

Response: We have now added this information to the Methods section and merged the two

Comment: 4) What is the method of analysis used (e.g. statistical analysis)? This is a requirement to be stated in the abstract since the term "significant" is used.

Response: We have added "questionnaires were analysed individually and using linear mixed-effect models".

Comment: 5) The Results section of the Abstract is very hard to understand. "Significant alignment suggests self-reported anxiety and 3, stable" is opaque in its meaning but can be used to understand, especially because the abstract often determines whether or not a reader will go on to read the whole article.

Is this sentence meant to mean "Self-reported anxiety and burnout, (patient-related burnout?), team understanding and work related burnout increased over time, all irrespective of alignment. Patient-related burnout, personal burnout, and work related burnout all increased, but were negatively related to Professional Identity. Males self-reported significantly lower Personal Burnout, but higher Patient Related Burnout, than females, but age and team understanding had no effect on burnout. Participants who undertook misaligned assistantships related higher patient- and work-related burnout."

However we also suggest that the last sentence "Participants who undertook misaligned assistantships reported significantly lower personal burnout" contradicts the first sentence of the conclusion "Participants experienced workplace learning that was either aligned or non-aligned to their first junior doctor post. An online questionnaire included: Professional Identity Scale, Team Understanding Scale, modified Hamilton Anxiety Rating Scale and modified Copenhagen Burnout Inventory. Data was collected on four occasions: (T1) prior to graduation; (T2) 1-month post-transition; (T3) 6-months post-transition; and (T4) 10-months post-transition. Questionnaires were analysed individually and using linear mixed-effect models".

Response: We apologize for the confusion. Firstly we analysed each scale by itself across time and examined whether there was an interaction between alignment/mis-alignment groups. Then we took the linear mixed effect model and examined the hypothesis that burnout is partly predicted by participants' experiences of assistantships alignment, professional identity and team understanding, with age and gender as confounding factors (thus comprised fixed effect components).

We have tired to specify these two different phases of our analysis in the Abstract as follows:

³$QDO\VHV RI LQGLYLGXDO VFDOHV UHYHDOHG WKDW VHOI UHSRUWH
related burnout were stable over time, while team understanding, personal and work-related burnout increased, all irrespective of alignment. Three linear mixed-effect models (with personal, patient-related and work-related burnout as outcome measures) found that males self-reported significantly lower personal, but higher patient-related burnout, than females, with age and team understanding having no effect. Anxiety was significantly positively related to burnout and professional identity was significantly negatively related. Participants experiencing non-aligned assistantships reported higher personal burnout and work- UHODWHG EXUQRXW OHYHOV RYHU WLPH ´

This way the reader can see that for the first analyses each outcome measure was examined separately ±but for the second analyses, the three levels of burnout were the outcome measure of the model.

Comment: 6) ³&RQFOXVLRQV 2XU ILQGLQJV EULQJ IRUWK LPSOLFDWLRQV
schools consider an end-of-year workplace alignment with first-post, or an extended shadowing period is considered immediately post-graduation. Educators might also consider how best to support XQGHUJUDGXDWH VWXGHQWV¶ HDUO\ 3, GHYHORSPHQW DQG VXSSRUW
transition for individuals with a predilection for anxiety. I think the authors wish, in the first sentence, to sa\ VRPHWKLQJ OLNH ³2XU ILQGLQJV VXJJHVW P end-of-year workplace DOLJQPHQW EHIRUH JUDGXDWLRQ RU DQ H[WHQGHG VKDGRZLQJ SHU VHFRQG VHQWHQFH , ZRXOG XVH SUHGLVSRVLWLRQ¶ UHGXFLQJ WKHLU positive desire!

5HVSRQVH :H KDYH QRZ DPHQGHG WKLV WR UHDG ³2XU ILQGLQJV EU
including that medical schools consider an end-of-year workplace alignment with first-post before graduation, or an extended shadowing period immediately post-graduation. Educators might also FRQVLGHU KRZ EHVW WR VXSSRUW XQGHUJUDGXDWH VWXGHQWV¶ HD
support systems should be in place across the transition for individuals with a predisposition for DQ[LHW\´

Comment: 7) 3DJH ³0HQWDO KHDOWK SUREOHPV DQG EXUQRXW REVHU
KDYH SRWHQWLDO UDPLILFDWLRQV´ , WKLQN WKH DXWKRUV PHDQ µLP

Response: Ramifications and implications are synonymous; we have now changed this terminology as the reviewer has a preference for this.

Comment: 8) 4:40 The professional socialisation of medical students has been of focus recently. µ:KDW GRHV µRI IRFXV¶ PHDQ"¶

Response: We have clarified this to rHDG ³KDV EHHQ WKH IRFXV RI PXFK VWXG\ UHF

Comment: 9) 4:43 develop a strong, sense of themselves as doctors. Stray comma.

Response: Comma deleted, thank you.

Comment: 10) ³VD-IUHHWC\DWHG LQFLGHQFHV´ , EHOLHYH WKH DXWKRUV P

Response: This has been amended.

Comment: 11) ³$GGLWLRQDOO\ WKHUH LV D JHQHUDO WHQGHQF\ IRU U DV ORZ DV Q DQG Q DW WLPH RQH´ 7KLV ZRXOG EH EHWWHU H[SU numbers are generally low H J Q DQG Q DW WLPH ´

Response: This has been changed

Comment: 12) 5:42 In some qualitative studies, participant numbers are also relatively low (i.e. n=15 and n=20). I would say that n=20 is low for a quantitative study, but may be excellent for a qualitative study, depending on the approach taken.

Response: We have removed this statement.

Comment: 13) µDQG LQ GLIIHUHQW FRQILJXUDWLRQV´

Response: Preposition has been added

Comment: 14) ³\RXU VWRU\´

Response: Corrected

Comment: 15) µPLVDOLJQHG DVVLVWDQWVKLSV¶

Response: We are unsure what this comment refers to ±there was no typo here.

Comment: 16) µPLVDOLJQHG¶ VXJJHVWV WKHUH LV DQ DFWLYH PLVPDWFK UHFRPPHQG WKHOLJQHG¶ RQH Accurate descriptor.

Response: This has now been done

Comment: 17) µD OLQHDZLDVHG¶

Response: Corrected

Comment: 18) 3DJH µ3URFHGXUH¶ 7KH VHFRQG DQG WKLUG SDUDJUDSKV paragraph to eliminate duplication.

Response: This has now been done and repetition has been removed.

Comment: 19) 7:7 Font size change

Response: This has been addressed

Comment: 20) µ$ WRWDO¶
21) 8:7 An total of n=281 (56% response rate aggregated over time) participated in this study, comprising: 68% (n=183) females and 73% (n=206) aged 17-25.
22) Here and elsewhere, the use of the colon seems unnecessary. This sentence works better as ³7KH UHVSRQVH UDWH DJJUHJDWHG RYHU WLPH ZDV Q RI ZKR 73% (n=206) were aged 17- ´

Response: This has now been changed to read: The response rate aggregated over time was 56% (n=281), of whom 68% (n=183) were female and 73% (n=206) were aged 17-25, as suggested.

Comment: 23) ³7KH SHUFHQWDJHV RI IHPDOHV E\ WLPH SKDVH«´

5HVSRQVH ³WKH´ KDV EHHQ RPLWWHG

Comment: 24) 7KXV WKH VL]H RI WKH LQWHUFHSWV ³
25) ³LV OLNHO\ WR EH GXH WR WKH PHDVXUHPHQW FKDUDFWHULVWL
this ±LVQ¶W LW GXH WR GLIIHUHQFHV LQ GLVGXO G

Response: Yes - We amend the description to state this more clearly as individual differences for burnout measures.

Comment: 26) ³7KH $$ SDUWLFLSDQW JURXS reported HPBurnout UHQFH JURXS V
increased and an average of 1.01 per WLPH SRLQW´ :KDW GRHV WKLV PHDQ"

Response: The score 1.01 is the slope for the fixed effect of Time. It is like beta in a regression line.
:H KDYH QRZ DGGHG D µQRWH¶ XQGHUQHDWK WKH)LJXUH WR WKLV H

Comment: 27) The authors may well be so familiar with their acronyms that they can decode them immediately. However, the reader is not in this position. Spelling out the acronyms will add slightly to the word length but very greatly to the clarity.

Response: This has now been done (as requested above)

Comment: 28) The authors report an effect of aligned versus non-aligned assistantships on the properties they measure. However, this is not under the control of the medical schools, Foundation programmes, or even entirely under the control of the graduates ±it results from the interaction between student preferences and their EPM and SJT scores in the Selection for Foundation process.
7KHUHIRUH WKHUH LV QRWKLQJ WKH V\VWHP FDQ FXUUHQWO\ GR WR
and the observations made in the paper are interesting but currently not useful. The authors propose an extended shadowing period post-graduation (presumably between the June exam and the beginning of August) but (a) there is no evidence that this will work ±it is purely speculative and (b) there are challenges to the practicality of this approach, given that it will require graduates to move to where they will be working in F1, before they are paid a salary.

Response: We tend to disagree with this comment ±at best it is fatalistic.

In terms of what medical schools can do: they can adjust the timings in their curriculum as has been done in Wales. We agree that this is both difficult and inconvenient but not impossible. Furthermore, ZH¶UH QRW WKH ILUVW SHRSOH WR FDOO IRU WKLV /LJKWPDQ (.LQJ
for final-year students. The Clinical Teacher 2015;12(2):115-20. doi: 10.1111/tct.12272). Furthermore, some final year students already undertake this informally with the support of their medical schools.

5HJDUGLQJ WKH DVVHUWLRQ WKDW ³7KHUH LV QR HYLGHQFH WKLV ZL
that there is a potential benefit and refute that this is purely speculative. Indeed, this study is set within a much larger body of research that comprises both qualitative and quantitative data to examine the assistantship programme (including: Jones, MJ, Okeke C, Bullock A, Wells SE,
0RQURX[H /9 ³+H¶V JRFLQRUWLRQ EHX JXYGW´ $ QDUUDWLYH LQWHUYL

VWXGHQWV¶ DQG WKHLU HGXFDWRUV¶ H[SHULHQFHV RI DOLJQHG DQG
6:e011817 doi:10.1136/bmjopen-2016-01181).

Comment: 29) In order to indicate the generalisability of the findings, it would be valuable if the authors indicated for the UK generally, how many graduates work locally to their medical school. This number varies greatly (for instance, London has a low match), and the data is available from UKFPO.

Response: We thank the reviewer for this suggestion and now include a comment on this aspect in our Discussion section:

[3]2XU ILQGLQJV EULQJ IRUWK VRPH LPSOLFDWLRQV IRU SUDFWLFH DQ
graduates experiencing an aligned assistantship demonstrate a lower rise in personal burnout over the transition period and a stabilized work-related burnout as compared with those whose assistantship was non- DOLJQHG ZH VXJJHVW WKDW PHGLFDO VFKR-ROV FRQVLG
year placements with their first-post on transition, or that an extended shadowing period is explored immediately on graduation. Although this suggestion might seem out of reach, we believe that it is not impossible. For example, in 2016 across the UK an average of 40.5% of graduating medical students selected their local foundation school (the organisation which provides postgraduate training) as their first preference for postgraduate studies (with Northern Ireland and Scotland having over 60% of graduates applying locally), the vast majority of which were successful.76 This means that potentially around 40% of UK graduates could experience an aligned assistantship prior to their transition into practice. Such a move would enable further research on the potential benefits of alignment to be H[DPLQHG RQ D PXFK ZLGHU VFDOH [3]

VERSION 2 – REVIEW

REVIEWER	Nancy Sturman Australia, University of Queensland
REVIEW RETURNED	04-Sep-2017

GENERAL COMMENTS	Thank you for addressing reviewer feedback, and clarifying sections of the manuscript. I think it would be important to include the items in the Team Understanding instrument (and indeed all the other instruments used) as an appendix, to enable the reader to assess the pertinence of these measures and the rather complex relationships between them.

REVIEWER	Bryan Burford Newcastle University, UK
REVIEW RETURNED	12-Oct-2017

GENERAL COMMENTS	Thank you to the authors for their detailed and considered responses

W R U H Y L H Z H U V ¶ F R P P H Q W V , K D Yhere to U H V W those areas noted in my earlier review. All are points of clarification of my earlier comments rather than indicating need for further revision, but may serve to sense-check.

I also found the provision R code helpful in understanding the analysis, but realise I may be an outlier in that.

Appendix B did not seem to be available to me.

Initial comment	$ X W K R U V ¶ Response	Clarification
A clear and comprehensive review. One slight quibble is that the implicit hypothesis that professional identity may vary with assistantship alignment is not clearly justified. As it is not stated as an explicit hypothesis that may not be too much of a problem, but if there is a theoretical rationale, it would be helpful to have it clearly presented.	We present this theoretical rational in the previous sections entitled: µ 0 R Y L Q J I U F individual to institutional factors I R U E X U Q R X µ F X U U L F X O D L Q W H U Y H Q W we set out the rationale for contextual factors affecting burnout and we also discuss other research that has attempted to examine the effect of assistantships on workplace stress. We believe that this information is sufficient to inform our hypothesis.	My comment related more to the data as presented in table 2. I had inferred there was a hypothesised difference between aligned and non-aligned participants, but think I may have been reading too much into the presentation of the data.
Under the analysis section, no details are given of model/variable selection ±I infer that no comparison of subsetted models was undertaken (to test for example the contribution of the interaction effects),	We did include the interaction effect of Group and Time in our model (see the formula in P7) and reported the interaction effect in Table 3 (in the column titled Group * Time of	Thank you for the clarification. I understand the non-inclusion of other interactions on theoretical and pragmatic groups. My query was also

	which as I understand the analysis is acceptable, but should probably be stated clearly. I found myself wondering about interactions between PI, anxiety and time, and it would be nice to know if these were considered or excluded.	P14). However, we did not do the interaction effect for subsetted models between time and other independent variables (i.e. PI or anxiety). We did not do this for two reasons: 1) the personal characteristics in PI or anxiety were taken as confounding variables for the effect of time and group which we were most interested in; 2) it would become extremely complicated due to the number of variables we have. To clarify this in our method, we added a sentence LQ 3 ³$OWK is interesting to see the interaction effects between these factors and time, we do not include these interaction terms in our model for simplicity SXUSRVH ´	about the approach to variable selection in general ±so whether any criterion-based, stepwise approach to model development was undertaken. It is clearer to me now that the model was theoretically rather than empirically determined.
	Details of drop-out	As you correctly	Just a clarification

	within each time point would be informative ± while not definitive, Bristol Online Surveys gives numbers of those who do not complete a questionnaire, and given the length of your form, it would be reassuring to know how few or many may have been deterred from completion by its length.	say, BOS will give out information regarding incomplete surveys. However, as it is designed to maintain respondent anonymity it will not automatically capture any identifying information about the respondents, we have to set up a respondent list apriori ±it can then tell us if a respondent started but did not complete the questionnaire. As we did not have access to potential S D U W L F L S D addresses before the study started (we contacted them via a third person) we did not set up the respondent list and so this information is not available to us.	of my point here ±I was referring to those who may µ U H W L U H ¶ S through completion at a given time point, rather than attrition through the study. So you can see how many (anonymous) people begin but do not complete the whole questionnaire, or if a lot of people exit on a particular page - W K D W G require a respondent list to be set up. , G R Q ¶ W W K changes anything substantive, so G R Q ¶ W W K L C adding, but provides a little info about how the questionnaire was approached by respondents.

VERSION 2 – AUTHOR RESPONSE

This is because the request by Reviewer 1 is:

³7KDQN \RX IRU DGGUHVVLQJ UHYLHZHU IHHGEDFN DQG FODULI\LQJ V
would be important to include the items in the Team Understanding instrument (and indeed all the other instruments used) as an appendix, to enable the reader to assess the pertinence of these PHDVXUHV DQG WKH UDWKHU FRPSOH[UHODWLRQVKLSV EHWZHHQ WK

±yet we already do this ±LW LV FOHDUO\ ODEHOOHG ³$SSHQGL[$´ DQG KDV EHH the start.

Reviewer 2 only made one request with the remaining of his comments thanking us for clarification. His request was as follows:

³-XVW D FODULILFDWLRQ ZRXOG EH SHUKDSV PHQWLRQ WKRVH ZKR PD\ µUHWLUH¶ completion at a given time point, rather than attrition through the study. So you can see how many (anonymous) people begin but do not complete the whole questionnaire, or if a lot of people exit on a particular page ±WKDW GRHVQ¶W UHTXLUH D UHVSRQGHQW OLVW WR EH VHV substantive, so doQ¶W WKLQN LW QHHGV DGGLQJ EXW SURYLGHV D OLWWOH DSSURDFKHG E\ UHVSRQGHQWV ´

- We have now added the following to the manuscript in order to clarify this:

³:H GLG QRW XVH ZLWKGUDZDO DV D 26 WR HQDEOH LQGLYLGXDOV WR µVD ODWHU¶ VR ZH DUH XQDEOH WR NQRZ KRZ PDQ\ SHRSOH VWDUWHG

We did not hear anything about our new Appendix B ±I think Reviewer 2 liked them so would you like us to keep them as an Appendix?

We trust you are now happy with our responses and I look forward to receiving your response.