Thank you for the opportunity to review the commentary entitled “The Sorting Hat: Cool Fiction Element but Not Necessarily a Good Career Advisor” on the paper which was itself entitled, “The Sorting Hat of Medicine: Why Hufflepuffs Wear Stethoscopes and Slytherins Carry Scalpels.” In reviewing the commentary, I also decided that reading the original paper (which I will henceforth refer to as “B-G&V” after the authors’ last initials) would be helpful. In writing this review, I tried to be mindful not to play the role of yet another reviewer of the B-G&V paper, as opposed to a reviewer of the commentary.

Overall, I agree strongly with one of the fundamental points raised in the commentary. In my opinion, the B-G&V did indeed ‘position’ the ‘sorting hat’ as an indicator of personality, which, in turn, might well be a predictor of specialty preference. I agree with the commenters that the connection between personality and Hogwarts house preference is far from clear, and that the data from B-G&V reveal nothing about that connection. In other words, personality seems to be a red herring in the original piece.

I would, however, like to see the commenters connect the argument above to their recommendations a little better. To make my point here clearer, let’s take the results of B-G&V at face value for the moment. If we do that, while yes, the connection with personality and the ‘sorting hat’ remains unclear, there still seems to be some sort of link between the sorting hat and the specialty that is chosen. Thus, on the surface at least, it would follow from the results that the sorting hat may well be a good tool for selecting a specialty, even if we had no idea why that was the case.

There are other reasons to be suspicious of this surface level conclusion, however, and perhaps the commenters could draw from these a little more. A first concern is that the analyses is conditional on awareness of the Harry Potter story, which itself might have a causal relation with specialty selection (e.g. if it was associated with higher openness or parental status). A second is that we have no information about the reliability of one’s stated preference (for all the reader knows, it could change from assessment to assessment), nor whether the experience in the chosen specialty has itself shaped the sorting hat result, after the fact. A third is that the B-G&V did not collect (or report at least) any data on satisfaction with or performance in the chosen specialty. Perhaps the key variable of interest here is merely specialty selection, but I would imagine that the ultimate purpose of helping people choose specialities is to help them select a specialty that they will be satisfied in and perform well in. Thus, while the ‘sorting hat’ might tell us who selects what in the absence of other details, it might not tell us about who thrives in what specialty. Without the latter connection, it would seem not to be a useful predictor.
Altogether, it was a very interesting experience to see how some of the ideas from my discipline are applied into others and I hope these comments are helpful!

Review 2
Completed June 5, 2020

Collabra Review –
The Sorting Hat: Cool Fiction Element but Not Necessarily a Good Career Advisor

In this commentary, the authors offer several loose critiques of a recently published article “The Sorting Hat of Medicine: Why Hufflepuffs Wear Stethoscopes and Slytherins Carry Scapelss.” I have now read both the original article and this commentary, and I certainly agree with the authors that the original article should be interpreted with great caution due to several substantial methodological flaws. In the interest of transparency, I admit that I struggled to determine my role as a reviewer here. Should I offer critiques of the original article, such that the current authors may consider appending them to their commentary? Or should I simply review this commentary as an independent piece? This review is ultimately a bit of both, and I hope that the editor and authors find these comments useful.

1) My primary comment is that I think the current authors could (and should) be much, much more critical of the original manuscript. The authors mention that “we do not seek to criticize Baimas-George and Vrochides’ research” – I don’t understand this. First, this isn’t entirely true as the authors do note a few criticisms (e.g., poor personality assessment), and second, I don’t really know what the purpose of this commentary is in the absence of such critical feedback. My recommendation is that the authors substantially revise this commentary to be explicitly critical of the flaws of the original article. In its current form, the commentary feels overly gentle and short of its potential.

2) If I were to write my own commentary of the original article, here a few points I’d be sure to include (in what would absolutely be a recommendation for Rejection):

2a) Although it is a somewhat unique sample, the use of a single item that participants used to self-sort into houses is an insurmountable flaw in my mind. As the authors note, there is “little empirical evidence” (pg. 1) to support the links between Harry Potter houses and personality traits/other individual differences. Thus, in the absence of evidence supporting the validity and reliability (e.g., convergent/discriminant validity with relevant, well-measured personality traits) of self-sorting into houses with a single item, I have no idea what to make of the results, and I’d argue they’re virtually uninterpretable in terms of linking personal characteristics to occupational decisions. With this in mind, I view most of the discussion section as wholly speculative and unfounded.
2b) The results in Table 2 are difficult to follow – I suggest making it more digestible. For example, it’s not immediately clear which comparisons the p-values are referencing, please clarify. Additionally, many of the results appear extremely under-powered. The authors make claims about differences in percentages of house endorsement when a given specialty (e.g., OB-GYN) may have less than 20, or even less than 10 total respondents. Again, even with a valid personality assessment, I’m not convinced that a sample of this size would permit any meaningful inferences. As a separate but equally important issue, despite conducting many tests of significance, there appears to be no consideration of multiple comparison correction.

2c) What the column “muggle” mean in this article? There is no reference to this term in any other part of the article, so I don’t how it was operationalized.

2d) Given the many fatal flaws detailed, I find it especially inappropriate that the original authors use the term “why” in the title. I’d argue they have presented very little scientific evidence of anything, much less do they approach the high bar of explaining “why” certain personality characteristics are related to particular occupational decisions.

3) Relatedly, in the abstract for this commentary, the authors make the point that “conclusions should not be drawn uncritically from statistically significant relations”, but discussion of this point is totally absent from the commentary. I suggest that the authors integrate this comment, and integrate a focus on effect size rather than statistical significance (see Funder & Ozer, 2019).


4) This is more stylistic than substantive, but I tend to find language like “spoil the magic by revealing the underlying trick” to be unbefitting of a scientific paper. Plus, what’s the “trick” that’s being revealed?

Editor decision submitted June 12, 2020  
Editor: Beth Visser  
Decision: Revise and resubmit

Dear Dr. Jarke,

Thank you for submitting this commentary to Collabra: Psychology. I enjoyed reading it, and you can see that both reviewers were also quite positive. I particularly appreciated that both reviewers read the Baimas-George and Vrochides (2020) article quite thoroughly in order to fully appreciate your commentary.
As you can see from the reviews, both were generally in agreement with your comments, but thought you could expand on them. Both had further critiques of the original article. I don’t ask that you incorporate all of these points or make drastic changes to your tone. However, I think that you could reflect on the comments and decide how you could strengthen your existing comments. For example, I thought Reviewer 1 offered some excellent potential explanations for a superficial connection between sorting hat determination of house and medical specialty, and I think you could incorporate some of this reasoning. Reviewer 2 pointed to some different issues (e.g., underpowered analyses) but offered a similar overall point that your commentary could go further. I won’t reiterate all of the reviewers’ points, but would ask that you read and reflect on them in thinking through a revised version of this commentary.

Thank you again for submitting this well-written manuscript to Collabra: Psychology.

Regards,

Beth Visser (PhD)

Revision submitted July 12, 2020
Author response to reviewers

Answers to Reviewer 1

Thank you for the opportunity to review the commentary entitled “The Sorting Hat: Cool Fiction Element but Not Necessarily a Good Career Advisor” on the paper which was itself entitled, “The Sorting Hat of Medicine: Why Hufflepuffs Wear Stethoscopes and Slytherins Carry Scalpels.” In reviewing the commentary, I also decided that reading the original paper (which I will henceforth refer to as “B-G&V” after the authors’ last initials) would be helpful. In writing this review, I tried to be mindful not to play the role of yet another reviewer of the B-G&V paper, as opposed to a reviewer of the commentary.

Overall, I agree strongly with one of the fundamental points raised in the commentary. In my opinion, the B-G&V did indeed ‘position’ the ‘sorting hat’ as an of indicator of personality, which, in turn, might well be a predictor of specialty preference. I agree with the commenters that the connection between personality and Hogwarts house preference is far from clear, and that the data from B-G&V reveal nothing about that connection. In other words, personality seems to be a red herring in the original piece.

We thank the reviewer for both, their comments, and for taking the time to go throughout the original article and our commentary. We understand the difficulty of conducting a review of a commentary such as the one presented, so we truly appreciate the reviewer for his/her work.
Following the reviewer's suggestions, we have delved on the different issues both, previously presented and raised now by their comments. For clarity’s sake, we proceed to comment on each specific issue detailed in the review.

I would, however, like to see the commenters connect the argument above to their recommendations a little better. To make my point here clearer, let’s take the results of B-G&V at face value for the moment. If we do that, while yes, the connection with personality and the ‘sorting hat’ remains unclear, there still seems to be some sort of link between the sorting hat and the speciality that is chosen. Thus, on the surface at least, it would follow from the results that the sorting hat may well be a good tool for selecting a speciality, even if we had no idea why that was the case.

There are other reasons to be suspicious of this surface level conclusion, however, and perhaps the commenters could draw from these a little more.

A first concern is that the analyses is conditional on awareness of the Harry Potter story, which itself might have a causal relation with specialty selection (e.g. if it was associated with higher openness or parental status).

This is an interesting point that we had not previously acknowledged in our discussion of B-G&V paper. As such, we believe it is worthy that readers are aware that familiarity with the Harry Potter world might have a direct effect on speciality selection (as discussed in the previous section of the text), but also an indirect effect as well via individual characteristics. Thus, we now present this additional limitation of their study. As such, we have included it in the text (page 3, first paragraph).

A second is that we have no information about the reliability of one’s stated preference (for all the reader knows, it could change from assessment to assessment), nor whether the experience in the chosen speciality has itself shaped the sorting hat result, after the fact.

We thank the reviewer for this comment. We agree that in order to pursue Harry Potter preference as a tool that would serve as a proxy of speciality selection, we need to understand whether it is reliable and valid or not. As occurs for any other psychological assessment tool, whether this information is missing, results and conclusions would be strongly compromised. And we agree that this information is again conspicuous by being absent in the original article. We have now provided an explicit account of this limitation in page 2, last paragraph.

A third is that the B-G&V did not collect (or report at least) any data on satisfaction with or performance in the chosen speciality. Perhaps the key variable of interest here is merely speciality selection, but I would imagine that the ultimate purpose of helping people choose specialities is to help them select a speciality that they will be satisfied in and perform well in. Thus, while the ‘sorting hat’ might tell us who selects what in the absence of other details, it
might not tell us about who thrives in what specialty. Without the latter connection, it would seem not to be a useful predictor.

We agree that one of the strong limitations of the original B-G&V proposals is that they limit their study to pure correlational methods without any psychometric information about neither the reliability nor the validity of their proposal. Furthermore, they also applied a strong casual language in certain sections of the text that is unwarranted given the methodology followed in this study. Following the reviewer suggestion, we have made explicit our strong concerns regarding the lack of information on these areas.

Altogether, it was a very interesting experience to see how some of the ideas from my discipline are applied into others and I hope these comments are helpful!

Again, we thank the reviewer for taking the time to read both, our commentary and the original article and for helping us to improve our article. We hope we have satisfactorily accounted all his/her comments.

Answer to Reviewer 2

The Sorting Hat: Cool Fiction Element but Not Necessarily a Good Career Advisor

In this commentary, the authors offer several loose critiques of a recently published article “The Sorting Hat of Medicine: Why Hufflepuffs Wear Stethoscopes and Slytherins Carry Scalpels.” I have now read both the original article and this commentary, and I certainly agree with the authors that the original article should be interpreted with great caution due to several substantial methodological flaws. In the interest of transparency, I admit that I struggled to determine my role as a reviewer here. Should I offer critiques of the original article, such that the current authors may consider appending them to their commentary? Or should I simply review this commentary as an independent piece? This review is ultimately a bit of both, and I hope that the editor and authors find these comments useful.

We thank the reviewer for their comments. We understand the difficulty of conducting a review of this kind of commentary, and we truly thank the reviewer for taking their time to go through our article and the original article. For clarity’s sake, and in order to address the different concerns raised by the review, we will address each of the comments provided separately.

1) My primary comment is that I think the current authors could (and should) be much, much more critical of the original manuscript. The authors mention that “we do not seek to criticize Baimas-George and Vrochides’ research” – I don’t understand this. First, this isn’t entirely true as the authors do note a few criticisms (e.g., poor personality assessment), and second, I don’t really know what the purpose of this commentary is in the absence of such critical feedback. My recommendation is that the authors substantially revise this commentary to be explicitly critical of the flaws of the original article. In its current form, the commentary feels overly gentle and short of its potential.
We thank the reviewer for this comment. We understand that presenting a critical review of a colleague’s article is an initiative that must be carefully and politely addressed. Energizing words of the reviewer, we have tried to present our concerns with a stronger voice and complemented by a more nuanced overview of our concerns.

Regarding the sentence: “we do not seek to criticize Baimas-George and Vrochides’ research”, we believe that it has been misunderstood due to poor writing on our end. We intend to reflect that what we understand that is the core idea of their article, using pop-culture references to create new tools for vocational purposes might not be ill-advised. It was just that the claims presented were untenable given their research approach and evidence provided in the paper. To avoid any further confusion, we have clarified this particular sentence in the text.

2) If I were to write my own commentary of the original article, here a few points I’d be sure to include (in what would absolutely be a recommendation for Rejection):

2a) Although it is a somewhat unique sample, the use of a single item that participants used to self-sort into houses is an insurmountable flaw in my mind. As the authors note, there is “little empirical evidence” (pg. 1) to support the links between Harry Potter houses and personality traits/other individual differences. Thus, in the absence of evidence supporting the validity and reliability (e.g., convergent/discriminant validity with relevant, well-measured personality traits) of self-sorting into houses with a single item, I have no idea what to make of the results, and I’d argue they’re virtually uninterpretable in terms of linking personal characteristics to occupational decisions. With this in mind, I view most of the discussion section as wholly speculative and unfounded.

We agree with the reviewer that the results presented are strongly compromised due the lack of information regarding the validity and reliability of measure used. Furthermore, the conclusions presented by the authors are unwarranted without providing this information. Following the reviewer suggestion, we have decided to present a more compelling and detailed argument regarding these limitations in the text (page 2-3, last/first paragraph).

2b) The results in Table 2 are difficult to follow – I suggest making it more digestible. For example, it’s not immediately clear which comparisons the p-values are referencing, please clarify. Additionally, many of the results appear extremely under-powered. The authors make claims about differences in percentages of house endorsement when a given specialty (e.g., OB-GYN) may have less than 20, or even less than 10 total respondents. Again, even with a valid personality assessment, I’m not convinced that a sample of this size would permit any meaningful inferences. As a separate but equally important issue, despite conducting many tests of significance, there appears to be no consideration of multiple comparison correction.
We thank the reviewer for this comment. We have included these methodological concerns in the text now (page 3, second paragraph).

2c) What the column “muggle” mean in this article? There is no reference to this term in any other part of the article, so I don’t how it was operationalized.

We agree with the reviewer that the original result section is, at times, presented in an incoherent fashion, with details such as the one highlighted by the reviewer, unclear. However, for our commentary we believe that is more beneficial for the readers if we comment on the big picture and the larger problems of the research presented and leave these other issues unmentioned. On a side-note: A muggle, in the Harry Potter franchise, refers to a person with no magical abilities and should have clearly been explained in the B-G&V paper.

2d) Given the many fatal flaws detailed, I find it especially inappropriate that the original authors use the term “why” in the title. I’d argue they have presented very little scientific evidence of anything, much less do they approach the high bar of explaining “why” certain personality characteristics are related to particular occupational decisions.

We agree that the original authors should have conducted a thorough exploration of the relationship between personality traits and career path decision-making, or at least to provide additional evidence on the underlying mechanisms connecting the House selection and personality traits. We have extended our criticism in the text, and presented altogether with a particular concern of the alternative reviewer, to better reflect these concerns.

3) Relatedly, in the abstract for this commentary, the authors make the point that “conclusions should not be drawn uncritically from statistically significant relations”, but discussion of this point is totally absent from the commentary. I suggest that the authors integrate this comment, and integrate a focus on effect size rather than statistical significance (see Funder & Ozer, 2019).

We thank the reviewer for this reference. Given its appropriateness, we have included in the section of our commentary when we address our concerns regarding the methodology followed by the original authors.

4) This is more stylistic than substantive, but I tend to find language like “spoil the magic by revealing the underlying trick” to be unbefitting of a scientific paper. Plus, what’s the “trick” that’s being revealed?

Following the reviewer comment, we have reviewed the text and changed a number of casual expressions used in the previous version. We have, however, retained some non-technical language as we believe it makes the paper more engaging without losing credibility.
Again, we thank the reviewer for taking the time to read both, our commentary and the original article and for helping us to improve our article. We hope we have satisfactorily accounted for all his/her comments.

Editor final decision
July 22, 2020
Accept

Dear Hannes Jarke,
I have now had a chance to read over your manuscript “The Sorting Hat: Cool Fiction Element but Not Necessarily a Good Career Advisor”, along with the letter describing the changes you made. Thank you for your responsiveness to the concerns that the reviewers and I raised. I am happy to say that your paper is now officially accepted for publication in Collabra: Psychology. Congratulations on this excellent work, I think it will make an important contribution to the literature and I look forward to seeing it published! I hope your experiences with Collabra: Psychology have been positive and that you will continue to consider it as an outlet for your work.

As there are no further reviewer revisions to make, you do not have to complete any tasks at this point. Our managing editor will contact you in case there are any pre-production file related questions. You will have an opportunity to check the page proofs before we publish your article. Thank you again for publishing in Collabra: Psychology.

Sincerely, Beth Visser