The authors propose to undertake a large online study across two separate samples to evaluate individual differences in inattentional blindness (IB). The first will look at cognitive measures (Ospan, rotation, TestMyBrain matrices test) and the second will look at personality scales (BFI-2 for big 5 personality, MPQ Absorption scale, ASRS - ADHD self-report scale, and the FFOCI - five factor obsessive-compulsive inventory specifically fastidiousness, perfectionism, and punctiliousness), with the goal of assessing whether individual differences predict differences in IB. The authors propose to use two IB tasks (“sustained” IB, “transient” IB) in both experiments so that they can assess differences in the most classic IB experimental conditions. This also allows them to examine IB in one task as a predictor of IB in another task.

I applaud the authors for the overall methods and design, which appear to be very robust and well thought out. The analysis plan is similarly well conceived. The authors have used their recent meta-analysis as primary motivation and for generating hypotheses for their proposed RR, which from my perspective is a robust starting position and justification of a study. Taking their meta-analytic findings at face value, and with an aim to collect 1000 subjects per experiment, this study will undoubtedly gain important insights and contribute well to the literature.

Below I have outlined some concerns I have, both major and minor. I consider my expertise to lay within attention and consciousness, so my comments and recommendations are largely drawn toward the IB aspect of the RR (not so much the other individual difference measures). Major concerns are mostly methodological and some conceptual. Minor are mostly to do with the presentation of the manuscript and some possible errors I noted.

Major

Measure of awareness and added robustness checks

The authors propose to use as their primary outcome measure a “liberal” criterion for classifying a subject as a noticer, and then add in two ‘robustness’ checks (“conservative” and “conservative+”). In principle I am very much in favor of this idea, but I mostly disagree with the way it is done here.

The proposed liberal criterion is as follows: simply categorize those that respond “yes, I saw something” as a noticer. With such a loose criterion, there is no identification/recognition of the unexpected object as ‘proof’ that the subject actually saw it. The experimenters are therefore saying they will take the subject at their word, rather than requiring evidence that they indeed had the perceptual experience in question. This is problematic for multiple reasons:

1. We know that differences in response criterion can impact and confound our measures in studies of consciousness (meaning that some subjects are more likely to say yes, even if they did not have the corresponding experience, and vice versa, which may undermine evidence for so-called ‘unconscious’ processing, see Yaron et al., 2024). One limitation of the standard IB task is that there is not much of a way of checking for this (but we are working on this currently). So, with this limitation in mind, it is generally better to take a more conservative measure, just to be sure.

2. With such a liberal criterion, it is left unclear how the additional data will be treated. For example, what if a subject says “yes, I saw a stimulus” but were incorrect in the forced choice? What would be done then? Will these subjects be treated as equal to those who were correct in the forced choice (all classified as noticers)? To me, that would be inappropriate. As another example, what if a subject says "yes" on both the critical trial and following divided attention trial, but is incorrect in the forced choice task on one and is correct on the other? From the current proposal, it seems these would be treated as equivalent, which is problematic.

There is both theoretical/principled reasons and empirical data to suggest non-trivial differences between different outcome measures of awareness, for example those that rely on mere detection (ie yes/no) compared with those that require some higher order identification/recognition of the stimulus (see Koivisto et al. 2017; Persuh, 2018). So I appreciate that a balance between these competing issues is needed, but I think that an IB study is probably the wrong place to examine this particular issue, and more to the point am left concerned that this ‘liberal’ metric is *too* liberal. This is particularly so since it is their primary measure (see page 19, line 38-40: "For all correlations of performance on this task with other measures and for attempts to predict noticing of unexpected objects from other measures, we used the liberal noticing criterion on the critical trial"). At its worst, then, this has the potential to undermine a great deal of the results. It could also even undermine the motivation behind the research, since this is motivated from previous studies that have looked at similar issues, and yet such a loose criterion is not common in these prior works (to my knowledge).

In addition to the above, can the authors elaborate on the choice of the “conservative+” criterion? Since no check of how subjects perform on the distractor task as a possible exclusion is proposed, I take it the rationale for this added robustness check is such an assessment. I thought it was odd to see it as a robustness check in the measure of noticing and am not confident this is adequate. Perhaps the authors can consider placing it more central to their pipeline, such as by using it in some manner as part of their exclusionary criteria? In a standard IB study this is quite critical because it rules out that, if a subject is inattentionally blind, it is not simply because they are not performing (or cannot perform) the task. I can see some rationale for not needing this in this proposed study since there are other cognitive tasks being undertaken (and so performance on these could in principle be used for a similar purpose), but what about experiment 2 which will only use personality measures? There is also the possibility that authors may wish to retain variance given the research question involves individual differences. Either way, I believe it is worth elaborating on/justifying further.

Absence of full attention trial

I’m curious as to why the authors chose not to include a full attention trial in either IB task? These are not always routinely used, but I do believe they are far more common than not, and there is evidence that its use (for excluding subjects) can influence results (see Hutchinson et al. 2022). To me, there is also sound principled

reason for its use as a manipulation check to ensure subjects are able to perceive the unexpected object under normal viewing conditions. It is of course unnecessary if noticing rates end up being at ceiling in the divided attention trial, but since we do not know this in advance, it is a rather essential (and easy to implement) check.

Characterization of load theory

The authors border on mischaracterizing load theory at points. The reason for this is that the load manipulation in this study is clearly a cognitive load one—in the low load group, subjects will be required to retain one tally and in the high load, two tallies. The authors later themselves state it is a cognitive load manipulation. I have a couple key problems:

1) They seem to selectively cite and/or miss key references at points. For example, page 21, line 15-17, which omits citing other research with different findings (De Fockert & Bremner, 2011), as well as two important meta-analyses conducted on load in IB (Hutchinson et al. 2022; Matias et al. 2022), that are clearly relevant here.
2) The authors state that cognitive demands can lead to a decline in noticing rates, but cognitive load may be better characterized as an inverted U, where maximal IB should occur when cognitive demands are optimally challenging. So under conditions with some standard degree of cognitive demand, an increase should lead to a *reduction* in IB—the increased cognitive demands lead to resources inadvertently spilling over to process task irrelevant information. Indeed this is also what leads to its distinctiveness from perceptual load. See De Fockert and Bremner 2011 for a characteristic example of this. Hutchinson et al. 2022 conclude that cognitive load is all over the place, so I appreciate that this issue may not be resolved, but this should not give the authors a free go ahead to simply ignore the issue and paint it however they choose, particularly since they state it is a theoretically relevant manipulation (page 16, line 22-24) and indeed are in a unique position where their RR may add important evidence toward this issue (even if it is not their priority).

Characterization of Inattentional Blindness

In my view, claiming that IB's "primary measure" is something that falls outside subjects intentions/attention, is theoretically presumptive and in some respects wrong. I refer specifically to page 6, lines 19-21: "the primary measure in an IB task is noticing of objects that explicitly fall outside the participants' intentions and attention". The reason is because multiple mechanisms underly IB. One of those mechanisms is well characterized by so-called “attention set”, since IB is less likely for unexpected objects that sync up with the subjects top-down attentional strategy. In these cases, noticing is boosted precisely because the unexpected objects falls within the subject’s ‘intentions and attention’, and so here this characterization is awkward at best.

My recommendation is to use terminology that is less presumptive and is more in line with descriptors or features of IB, for example task relevance and expectations—the

unexpected object is task irrelevant in that there is no explicit response required of it to perform the task, and it is unexpected in that the subject has no prior knowledge it will be presented. IB can be spoken of using these terms without making claims to what it is a primary measure of (which may be more prone to bias based upon an authors preferred theory, see Yaron et al. 2022).

Stimulus size

The authors propose to undertake their work online which is a great opportunity to allow for a large sample size, something which is especially useful for IB. But I was surprised that they do not appear to propose to use any control for visual angle. This may not have been possible previously in online behavioural experiments, but is now routinely possible. One potential option that exists is to add in a virtual chinrest (Li et al. 2020). I am guessing the authors are apt with javascript since this is how their IB tasks are built. There may still be some noise in the measurements but it is much improved compared with nothing. I recommend the authors add in a virtual chinrest or justify why one (or something similar) has not been used here.

Minor

page 4, lines 24-27: the meta-analysis by Hutchinson et al. (2022) looks at lots of these ‘systematic’ factors, so would be worth mentioning

page 8, line 9-13: I am not entirely sure if this is the most common; perhaps consider revising wording to “one of the most common”

page 13, line 24: “There actually was an extra object”. I checked both tasks and this is not in the instruction on the transient task. The sustained task also seems to cease after about 1s for me. It could be an issue with my internet/web browser, or the demo version, but I tried on several computers and each time it occurred. It could be worth rechecking.

page 16, line 19-21: and whether the UO went from left-right or right-left? and the UO itself? Probably worth including all randomization factors here.

page 19, line 37-38: It is minor, but I recommend including in the main manuscript

page 20, line 8: There is a subheading "sustained IB" and then immediately following: "as for the transient IB task". I get what this is saying, but it is confusing. I suggest simple rephrase.

page 22, line 15-19: one potential problem is that, unlike many previous studies, only two pre-critical trials are proposed to be used. This might be insufficient for estimating accuracy (in such a way as to make it comparable to other studies, which seems to be the goal). Other studies using more pre-critical trials will mean greater precision for estimating subjects accuracy.

page 23, line 15: (xx); is this a missing citation?

page 34, line 22: bit unclear on why this is described as an exploratory analysis in the RR stage 1. My understanding is that this terminology should be reserved for stage 2 analyses/results that are not pre-registered.

Another question - was the unexpected object the same across trials? (so if it is an L for critical trial, it is an L for divided attention) It is a bit unclear at present.

References

de Fockert, J. W., & Bremner, A. J. (2011). Release of inattentional blindness by high working memory load: Elucidating the relationship between working memory and selective attention. Cognition, 121(3), 400–408. https://doi.org/10.1016/j.cognition.2011.08.016

Hutchinson, B. T., Pammer, K., Bandara, K., & Jack, B. N. (2022). A tale of two theories: A meta-analysis of the attention set and load theories of inattentional blindness. Psychological bulletin, 148(5-6), 370-396. https://doi.org/10.1037/bul0000371

Koivisto M, Grassini S, Salminen-Vaparanta N, & Revonsuo A. (2017). Different electrophysiological correlates of visual awareness for detection and identification. Journal of Cognitive Neuroscience, 29(9):1621–1631. doi: 10.1162/jocn_a_01149

Li, Q., Joo, S. J., Yeatman, J. D., & Reinecke, K. (2020). Controlling for Participants' Viewing Distance in Large-Scale, Psychophysical Online Experiments Using a Virtual Chinrest. Scientific Reports, 10(1), 1-11. doi: 10.1038/s41598-019-57204-1

Matias, J., Belletier, C., Izaute, M., Lutz, M., & Silvert, L. (2022). The role of perceptual and cognitive load on inattentional blindness: A systematic review and three meta-analyses. Quarterly journal of experimental psychology, 75(10), 1844–1875. https://doi.org/10.1177/17470218211064903

Persuh, M. (2018). Measuring Perceptual Consciousness. Frontiers in psychology, 8, 2320. https://doi.org/10.3389/fpsyg.2017.02320

Yaron, I., Melloni, L., Pitts, M., & Mudrik, L. (2022). The ConTraSt database for analysing and comparing empirical studies of consciousness theories. Nature Human Behaviour, 6(4), 593–604. https://doi.org/10.1038/s41562-021-01284-5

Yaron, I., Zeevi, Y., Korisky, U., Marshall, W., & Mudrik, L. (2024). Progressing, not regressing: A possible solution to the problem of regression to the mean in unconscious processing studies. Psychonomic bulletin & review, 31(1), 49–64. https://doi.org/10.3758/s13423-023-02326-x