Reply to Fleischman and Fessler's (2018) comment on "Hormonal Correlates of Pathogen Disgust: Testing the Compensatory Prophylaxis Hypothesis" (in press at Evolution & Human Behavior)

Benedict C Jones[1], Amanda C Hahn[2], Claire I Fisher[1], Hongyi Wang[3], Michal Kandrik[4,5], Anthony J Lee[1], Joshua M Tybur[4,5], Lisa M DeBruine[1]

1. Institute of Neuroscience & Psychology, University of Glasgow, UK.
2. Department of Psychology, Humboldt State University, USA.
3. School of Psychology and Cognitive Science, East China Normal University, China.
4. Department of Experimental and Applied Psychology, VU Amsterdam, Netherlands.
5. Institute Brain and Behavior Amsterdam, Netherlands.

Corresponding author: Benedict C Jones (ben.jones@glasgow.ac.uk), Institute of Neuroscience & Psychology, University of Glasgow, UK.

This research was supported by ERC Grants awarded to BCJ (OCMATE), LMD (KINSHIP), and JMT (HBIS).

We thank Fleischman and Fessler (2018) for their thoughtful and constructive comments on our paper "Hormonal Correlates of Pathogen Disgust: Testing the Compensatory Prophylaxis Hypothesis". We agree that improving measures of both disgust sensitivity and immunocompetence may yet reveal evidence for the Compensatory Prophylaxis Hypothesis and can only strengthen work on this topic. We elaborate here on these issues and suggest some directions for future research and methodological improvements.

We fully agree with Fleischman and Fessler (2018) that reliance on self-report measures of disgust sensitivity is an important limitation of our study (and previous studies on this and related topics). As we noted in Jones et al. (2018a), and Fleischman and Fessler (2018) reiterated in their comment, self-report instruments may not be optimal for testing the Compensatory

Prophylaxis Hypothesis, since they might not be able to detect small changes in disgust sensitivity. We proposed that facial electromyography, which has been used previously to assess inbreeding avoidance (De Smet et al., 2014) and disgust conditioning (Borg et al., 2016), could be used to test the Compensatory Prophylaxis Hypothesis. Other methods that are not dependent on self report, but that have not yet been used to investigate pathogen disgust (e.g., mouse-tracking paradigms, Freeman, 2018, and key-press tasks, Aharon et al., 2001), could also be adapted to test the Compensatory Prophylaxis Hypothesis. We reiterate our belief that such measures could yet reveal evidence for the Compensatory Prophylaxis Hypothesis not apparent in studies using self-report measures.

We also fully agree with Fleischman and Fessler's (2018) observation that progesterone might not fully tap the type of immunomodulation that sits at the core of the Compensatory Prophylaxis Hypothesis. Fleischman and Fessler (2018) refer to studies that have measured immunocompetence in other ways, noting that these measures of immunocompetence predicted avoidance of infectious disease cues. For example, they cite Miller and Maner (2011), who operationalized immunocompetence by dividing participants into "recently sick" (based on reporting having a head cold in the last two weeks) and "not recently sick" (based on reporting not having a head cold in the last two weeks). This approach might well bear fruit, but it awaits further validation, much like many other methods in the behavioral immune system literature (Tybur et al., 2014). Thus, we agree that alternative operationalizations of both immunocompetence and disgust could yet reveal compelling evidence for the Compensatory Prophylaxis Hypothesis.

Fleischman and Fessler usefully summarize a small literature evaluating the Compensatory Prophylaxis Hypothesis. We would like to take this opportunity to draw further attention to the importance of statistical power and valid measurement, not only in this literature, but throughout the evolutionary behavioral sciences. As many researchers have now noted, studies of hormonal regulation of behaviors and perceptions have typically been underpowered (e.g., Gangestad et a., 2016; Jones et al., 2018b). Previous

tests of the Compensatory Prophylaxis Hypothesis also faced this problem (Fleischman & Fessler, 2011; Zelazniewicz et al., 2016). Table 1 shows the power each of these studies had to detect each of the effects they reported. We modestly suggest that future studies in this literature aim to use the methods described in our paper, which combined a large sample size (i.e., number of women), multiple observations across the menstrual cycle, multilevel modeling, and salivary hormone measures to increase power. While we acknowledge that hormone measurement is costly and that other practical considerations are non-trivial, we note here that studies testing for effects of cycle phase on behavior are able to achieve high power and robust effects in the absence of hormone measurements by testing large numbers of women multiple times (see, e.g., Arslan et al., 2017 for a recent example of a high-powered test for fertility-linked changes in behavior with robust results that did not use hormone measures). Addressing this issue of power is a straightforward way to improve the replicability of work in this area.

Study	**Dependent variable**	**N**	**Reported effect (r)**	**Power to detect reported effect**
Fleischman & Fessler (2011)	contamination and washing compulsion	79	.31	.80
	bathroom behaviors	56	.29	.59
	disgust ratings of pictures	79	.25	.61
	ectoparasite-grooming	76	.25	.59
Zelazniewicz et al. (2016)	ds-r score	30	.41	.63
	core disgust	30	.29	.34
	animal disgust	30	.42	.65
	contamination disgust	30	.40	.61
	pathogen disgust	30	.40	.61
	moral disgust	30	.08	.07

Table 1. Power to detect effects reported in Fleischman and Fessler (2011) and Zelazniewicz et al. (2016). Only effects for luteal phase progesterone are shown from Zelazniewicz et al. (2016). See https://osf.io/93n2d/ for details of how power was calculated for these effects.

Finally, we would like to take this opportunity to highlight a second set of null results, which also speak to issues of statistical power and valid measurement in the evolutionary behavioral sciences. Sexual disgust has been proposed as

functioning to motivate avoidance of behaviors that would compromise an individual's reproductive fitness (Fessler & Navarrete, 2003). Researchers have hypothesized that such behaviors track changes in hormonal status (i.e., avoidance of behaviors that would compromise an individual's reproductive fitness will be particularly pronounced when fertility is high). We are aware of only one study to report a test of this hypothesis.

Using a cross-sectional design of 307 normally cycling women, Fessler and Navarrete (2003) reported that conception risk (assessed by a forward-counting method) was positively associated with a self-report measure of sexual disgust. According to estimates by Gangestad et al. (2016), the forward-counting method used in this study only correlates with actual conception risk at $r = .52$, and over 1000 participants are required to achieve 80% power to detect an effect size of the magnitude reported by Fessler and Navarrete (2003). Further, the face validity of the four-item measure of sexual disgust employed by Fessler and Navarrete is arguably suspect. It includes items such as "I think homosexual activities are immoral" and "I think it is immoral for people to seek sexual pleasure from animals" (Haidt et al., 1994). To the best of our knowledge, ours is the first published study since Fessler and Navarrete's to test this hypothesis, and it is the first to test how hormone levels relate to sexual disgust. We believe that (1) our design allowed much higher statistical power to detect small effects, and (2) our measure of sexual disgust, which includes items related to sexual choice (e.g., "A stranger of the opposite sex intentionally rubbing your thigh in an elevator", Tybur et al., 2009), provides a more valid assessment of the construct specified by Fessler and Navarrete. With these improvements, we found no evidence that sexual disgust tracked changes in women's hormonal status. That said, Fessler and Navarrete's hypothesis is logically compelling, and we do not suggest rejecting it based purely on our null result. Instead, we emphasize the importance of further improvements in study design and measurement. We suggest these improvements will move both this research area and the broader evolutionary behavioral sciences forward.

References

Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E., & Breiter, H. C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. Neuron, 32, 537-551.

Arslan, R. C., Schilling, K. M., Gerlach, T. M., & Penke, L. (2017). Using 26 thousand diary entries to show ovulatory changes in sexual desire and behaviour. PsyArXiv. https://psyarxiv.com/jp2ym/

Borg, C., Bosman, R. C., Engelhard, I., Olatunji, B. O., & de Jong, P. J. (2016). Is disgust sensitive to classical conditioning as indexed by facial electromyography and behavioural responses? Cognition & Emotion, 30, 669-686.

De Smet, D., Van Speybroeck, L., & Verplaetse, J. (2014). The Westermarck effect revisited: a psychophysiological study of sibling incest aversion in young female adults. Evolution & Human Behavior, 35, 34-42.

Fessler, D. M., & Navarrete, C. D. (2003). Domain-specific variation in disgust sensitivity across the menstrual cycle. Evolution & Human Behavior, 24, 406-417.

Fleischman, D. S. & Fessler, D. M. T. (2018). Response to “Hormonal Correlates of Pathogen Disgust: Testing the Compensatory Prophylaxis Hypothesis”. Evolution & Human Behavior, in press.

Fleischman, D. S., & Fessler, D. M. (2011). Progesterone's effects on the psychology of disease avoidance: Support for the compensatory behavioral prophylaxis hypothesis. Hormones & Behavior, 59, 271-275.

Freeman, J. B. (2018). Doing psychological research by hand. Current Directions in Psychological Science, in press.

Gangestad, S. W., Haselton, M. G., Welling, L. L., Gildersleeve, K., Pillsworth, E. G., Burriss, R. P., ... & Puts, D. A. (2016). How valid are assessments of conception probability in ovulatory cycle research? Evaluations, recommendations, and theoretical implications. Evolution & Human Behavior, 37, 85-96.

Haidt, J., McCauley, C., & Rozin, P. (1994). Individual differences in sensitivity to disgust: A scale sampling seven domains of disgust elicitors. Personality & Individual differences, 16, 701-713.

Jones, B. C., Hahn, A. C., Fisher, C., Wang, H., Kandrik, M., Lee, A., Tybur, J. M. & DeBruine, L. M. (2018a). Hormonal correlates of pathogen

disgust: Testing the Compensatory Prophylaxis Hypothesis. Evolution and Human Behavior, 39, 166-169.

Jones, B. C., Hahn, A. C., Fisher, C., Wang, H., Kandrik, M. & DeBruine, L. M. (2018b). General sexual desire, but not desire for uncommitted sexual relationships, tracks changes in women's hormonal status. Psychoneuroendocrinology, in press.

Miller, S. L., & Maner, J. K. (2011). Sick body, vigilant mind the biological immune system activates the behavioral immune system. Psychological Science, 22, 1467–1471.

Tybur, J. M., Frankenhuis, W. E., & Pollet, T. V. (2014). Behavioral immune system methods: Surveying the present to shape the future. Evolutionary Behavioral Sciences, 8, 274.

Tybur, J. M., Lieberman, D., & Griskevicius, V. (2009). Microbes, mating, and morality: individual differences in three functional domains of disgust. Journal of Personality & Social Psychology, 97, 103.

Żelaźniewicz, A., Borkowska, B., Nowak, J., & Pawłowski, B. (2016). The progesterone level, leukocyte count and disgust sensitivity across the menstrual cycle. Physiology & Behavior, 161, 60-65.