Assignment 2

Hunter Sterk; Dirk Keller; Markos Polos

Problems in task design.

(1) Stimuli and mask presentation. Gomez and colleagues (2016) utilized red-blue anaglyph glasses to present target images and masks to different eyes. However, stimulus presentation through anaglyphs glasses is prone to stereoscopic cross transference and, instead, a mirror stereoscope should have been used to guarantee separate stimulation of the two eyes (as it was done by Gayet & Stein, 2019; Baker, Kaestner, & Gouws, 2015).

(2) Assessment of ocular dominance. With respect to the assessment of ocular dominance, Miles's test (a hole-in-the-card test) does only yield binary result (dominant or nondominant) and does not take the interindividual variability of dominance into account, which might, as potential confound, compromise the b-CFS masking procedure. In addition, it is not recommended to use tests measuring sighting eye dominance to determine the dominant eye in a subsequent experiment involving interocular conflict. Both methods seem to rely on different suppression mechanism and hence behavioral outcomes of both methods are (often) not correlated. Rather, pretrials of the same task that will be used in the main experiment should be used for the priori quantification of eye dominance to ensure consistency and generalizability for ocular dominance between the assessment and experiment (Ding, Naber, Gayet, Van der Stigchel, Paffen, 2018).

(3) Control for emotional load of stimulus categories. Although Gomez and colleagues (2016) controlled for differences in emotional valance and arousal between image categories, unlike in Gayet and Stein (2019), the rating of emotional load did not take place under comparable experimental circumstances (or not mentioned otherwise). For example, rating emotional content of the stimuli could be assessed by flashing the stimulus on the screen with a short time interval in the same experimental set up. Given the premises of Gomez and colleagues (2016) that the impression on the retina elicit an emotional response, is fast and (nearly) subconscious, prolonged rating time of the stimulus sets (~6s as stated by the SEM protocol), however, does not reflect unconscious processing and allows recruitment of higher cognitive processes (episodic memories, expectations, etc.). This, again, compromises the generalizability between assessment and experimental execution of the emotional load between the stimulus categories. In addition, the inferences drawn from the differences in reaction times on the detection task, by Gomez and colleagues (2016), are argued to involve a fear guided action system. The emotional measures of valance (pleasant-unpleasant) and arousal (aroused-calm) are, however, a rather coarse measurements for fear and are not suitable to extract or isolate the subjective fear load of the stimuli. Important to note is that other emotions elicited by the target stimuli, such as disgust or happiness, do also have impact ratings on valance and arousal as well. To address this specificity, other research (Stein, Seymour & Hebart, 2014) directly investigated fear impression, rather than arousal and valance.

(4) Between-subject design. Although, a between-subject design was used for the a priori assessed ocular dominance of each subject, the stimuli were always presented to the right eye and the mask to the left eye of the participant, neglecting whether the participant had an ocular dominance for the right or left eye. Stimuli and mask should be presented in both dominant and nondominant eye (difference in right and left eye dominance in processing visual stimuli) since there might be an effect of dominance for the mask and the stimulus to the differences in reaction times. Moreover, a pure within subject design might have been the better choice (to avoid sample effects of a between subject design), such that the stimuli and mask could be consistently presented to the nondominant or dominant eye respectively (depending on the participants side of eye dominance).

(5) Controlling other methodological confounds. Regarding the set-up of the experiment, no gamma-correction of the monitor, lightening conditions of the room nor heads stabilization (chin- and headrest) were mentioned. Altogether confounding variables that could affect the measurement. In addition, no information was provided about the data-analysis procedure itself, such as software and packages, or any hardware specification, except for the model of the monitor.

(6) Reproducibility. The absent of relevant details regarding the equipment, methodology and analysis has serious adverse implications for the reproducibility of their results. In a similar study, by Grassini, Holm, Railo and Koivisto (2016), it was explored whether images of snakes provoke more fear due to evolutionary reasons. Colored images of snakes, spiders, birds and butterflies were presented and behavioral (subjective judgments and EEG data) were collected. The results showed that snake images had no advantage in breaking through consciousness in any of the masking conditions, nor were participants able to better distinguish snakes from other animals, either consciously or unconsciously.

Problems related to the stimuli.

(1) Low-level visual stimuli confound(s). Although it is known that low-level visual differences between image categories can and do often partially explaining difference in reaction time measures, particularly in the b-CFS paradigm (Ding, Naber, Gayet, Van der Stigchel, & Paffen, 2018), other low-level visual features besides spatial frequencies and mean luminance, such as curvature, elongation, or perimeter-to-surface ratio, are relevant as well and were not considered in the present study.

(2) Perimeter-to-surface ratio. With respect to the perimeter-to-surface ratio, Gayet and colleagues (2018) demonstrated that, "upon visual inspection of the 'snake' and 'bird' image categories (used in the study by Gomes et al., in press), […] it was apparent that the snakes in the images comprised more edges that bordered the uniform image background than the birds (or: a higher perimeter-to-surface ratio)". Difference in perimeter-to-surface-ratio between image categories (bikes and cars) were proven to affect mean reaction times in image detections (Gayet et al., 2018).

(3) Grating orientation and mean luminance. With respect to other low level image properties – besides the differences in perimeter-to-surface-ratio – the relationship of gradient orientation (HOG) and reaction time was not considered and, although, mean luminance across the stimulus categories did not differ significantly, the variance in luminance within the stimulus category is quite high for the snake images. On face-value, single snake images seem to strongly deviate from the adjusted, averaged background luminance. A high within-category variance for snake images might affect detection performance, since very high or low luminant target stimuli might breach through the mask facilitating faster detection of the target stimulus. More importantly, unlike in a comparable experiment conducted by Stein and colleagues (2014) the stimuli were not normalized for luminance and Root-Mean-Squared contrast, compromising standardization of the stimulus and its comparability regarding these low-level visual features across conditions.

(4) Contrast. Differences in contrast were not considered at all, even though it is known that the sensitivity of the visual system is not only affected by contrasts but does contrast also neuronally limit the perceivability of the stimulus (see contrast sensitivity function; Campbell & Green, 1965). Again, contrast should be normalized or equalized across conditions and stimulus background. Thus, assessing an images spatial frequency is not sufficient, but the interaction of both contrast and spatial frequencies determines the perceivability of a visual input.

(5) Ecological validity: Gray scale. The usage of gray scale images provides the advantage of eliminating RGB-related confounding effects, but, on the other side, compromise the ecological validity, thus its generalizability to real-life naturalistic scenes. This is especially problematic, since the authors attempt to not only generalize the results to predator-prey interactions in nature or naturalistic environments, but also interspecies similarities, evolutionary processes and neurological substrates that often have a strong relationship to color processing (also high spatial frequency encoding is predominantly found in cone-photoreceptors relevant for colour processing; Wolfe et al.,2006). In addition, grey scale images might be problematic because poisonous animals are often colored to avoid physical conflict, hence color is directly related to the stimulus category (snakes).

(6) Ecological validity: Motion. Spiders and snakes use camouflage to hide in cluttered scenes, as mentioned by authors, therefore making motion detection an essential, if not the most crucial, factor to detect camouflaged predators. Yet, the stimulus material used in the experiment is static. To allow generalizability to predator-prey events in nature, naturalistic video recordings would be an improvement on static grey scale images that do not allow drawing such strong conclusions for naturalistic phenomena.

(7) Degrees of visual angle. Lastly the spatial frequency analysis should be performed in cycles per degree of visual angle (not per pixel), because the image projected on the retina is concavely curved and pixel representations ultimately do not apply here anymore. Furthermore, the authors are inconsistent about these units within their report; low-level analysis uses pixel, while the stimulus is described in degrees of visual angle.

Problems in the analysis.

(1) Statistical assumptions, transformations and other corrections. Regarding the analysis, the report does not mention anything about a possible (non-)violation of the assumptions of statistical tests. In particular, normality might be a problem for the ANOVA, since reaction times are naturally skewed, which can seriously impact the significance test (or value). In addition, if the assumption of normality was, or must be, violated, which transformations or whether and what for transformation (normalizing, scaling, log-transformations) were performed to normalize the distribution. Secondly, no correction was applied to ensure that differences in reaction times do not reflect differences in personal response criteria (liberal versus conservative). Thirdly, trial attrition is quite high (<10%), as compared to Stein and colleague's (2014) original procedure (<2%). Neither a detailed description about the pattern of no- or wrong responses was provided, nor did they perform an analysis of the high trial attrition, such as comparing the proportion of targets that were (correctly) reported within the 7-second response deadline between the image categories (Gayet et al., 2019). This is important, because removing an unequal number of responses for each condition could introduce bias in further analysis. Fourthly, no power analysis was included in the paper that estimates the probability of finding a true positive result, given the sample and effect size.

(2) Aggregated data. In the present study only aggregated data was used and an analysis of individual difference in detection time performance across conditions were not conducted. Several standards confounds, such as simple demographics (age, gender, educational level), optical or emotional (fearfulness, phobias) of the participant can affect the differences between conditions.

Additional issues.

(1) Soundness of arguments. Perhaps, one of the most important criticism address their structure of argumentation. Importantly, the authors conduct a behavioral experiment and collect behavioral data. However, from these data (differences in reaction times to stimulus detection) inference about neurological properties, such as accessibility to the visual systems, are drawn. Such a conclusion, however, is beyond theoretical and technical limits of behavioral measurement and the interpretation of their results is not valid for many reasons: Most importantly, between input to the visual cortex and the motor execution lie multiple neuronal functions that are relevant for the whole process of reaction to a visual stimulus, such as low-level visual processes (contrast, acuity, purkinje shift) as well as high level visual processes (object recognition, space perception, depth analysis), information processing speed, memory, decision-making processes, motor planning and coordination and possibly more. All those functions are part of the process involved in the reaction time measure and any difference in those processes will result in different reaction times (e.g., differences in low-level or high-level visual processing) are relevant for reaction time. In addition, to assess their research interest on this level of granularity (accessibility to the visual system), neuronal recordings, such as single cell recording (patch clamping), should rather be used, because accessibility addresses mainly neurological properties of s specific neuronal circuit, such as early visual processing or a dedicated colliculus-pulvinar-amygdala pathway involved in 'fast' fear processing of visual content.

(1) Evolutionary psychology. Evolutionary psychology is speculation and its true nature, the cognitive states or mind of a species ancestor can never be investigated or falsified. It is therefore important to mention that the conclusions on an evolutionary level, drawn in cognitive or neurobiological experiments, are 'supportive' assumptions at best and no evidence for any evolutionary processes or neurobiological/cognitive properties evolutionary ancestors.

Summary.

Several concerns were discovered about the paper by Gomez and colleagues (2016). **(1)** With regard to the task design, the assessment of ocular dominance and emotional load of the stimuli are not done properly and can compromise its generalizability to the experimental procedure. In addition, anaglyph glasses are not the proper tool to present mask and stimuli in isolation to both eyes (a mirror stereoscope could be used instead). Moreover, the between subject design might be prone to left-right eye ocular dominance difference. A pure within subject design might be the better choice. Furthermore, other methodological aspects relevant for reproducibility of confound control, such as monitor gamma-correction, room lightening were not mentioned at best. The lack of experimental control has serious implication for the reproducibility of the experiment and other experiments have failed to reproduce their results. **(2)** With respect to the stimulus design, several low-level visual properties were not properly accounted for. For example, contrast and grating orientation were not assessed, perimeter-to-surface ratio differed between stimuli were not considered, high variability in luminance within a stimulus category exist and grey scale and motion absence in the experiment might compromise its generalizability to a naturalistic setting (since both are relevant to the theoretical context of snake threat detection). Since, images of bird and snakes are inherently different in terms of their visual characteristics, it remains uncertain whether a faster detection of snakes reflects a processing advantage for threat-unrelated visual properties, or a processing advantage for threatening stimuli. **(3)** Regarding the analysis, premises of the statistical tests were not mentioned or not checked. Moreover, no correction was applied to ensure differences in reaction times do not reflect differences in personal response criteria, trial attrition was quite high (<10%), as compared to other papers and no power analysis was included. Besides that, only aggregated data was used and several standard confounds could affect the differences between conditions. **(4)** Lastly, interpreting the found differences in detection times in the light of faster accessibility to the visual system as well as its implication in any evolutionary process or adaptation to selection pressure is questionable.

References:

Baker, D. H., Kaestner, M., & Gouws, A. (2015, January). Comparison of crosstalk for several methods of stereoscopic presentation. In *Perception* (Vol. 44, No. 4, pp. 455-455). London, England: PION LTD.

Campbell, F. W., & Green, D. G. (1965). Optical and retinal factors affecting visual resolution. *The Journal of physiology*, *181*(3), 576-593.

Ding, Y., Naber, M., Gayet, S., Van der Stigchel, S., & Paffen, C. L. (2018). Assessing the generalizability of eye dominance across binocular rivalry, onset rivalry, and continuous flash suppression. *Journal of Vision*, *18*(6), 6-6.

Gayet, S., Stein, T., & Peelen, M. V. (2019). The danger of interpreting detection differences between image categories: A brief comment on "Mind the snake: Fear detection relies on low spatial frequencies" (Gomes, Soares, Silva, & Silva, 2018). Retrieved from ResearchGate.com: https://www.researchgate.net/publication/331109585

Gomes, N., Silva, S., Silva, C. F., & Soares, S. C. (2017). Beware the serpent: the advantage of ecologically-relevant stimuli in accessing visual awareness. *Evolution and Human Behavior*, *38*(2), 227-234.

Grassini, S., Holm, S., Railo, H. & Koivisto, M. (2016). Who is afraid of the invisible snake? Subjective visual awareness modulates posterior brain activity for evolutionarily threatening stimuli. *Biological Psychology, 121, (2016) 53 –61*

Stein, T., Seymour, K., Hebart, M. N., & Sterzer, P. (2014). Rapid fear detection relies on high spatial frequencies. *Psychological science*, *25*(2), 566-574.

Wolfe, J. M., Kluender, K. R., Levi, D. M., Bartoshuk, L. M., Herz, R. S., Klatzky, R. L., & Merfeld, D. M. (2006). *Sensation & perception* (pp. 242-245). Sunderland, MA: Sinauer.